

Science needs the commission, *Sturm und Drang* and all

Despite its burdensome traditions, the European Commission may be the best hope for funding Europe's basic research.

Blood was metaphorically spilled in a small room in Brussels last week, when a dozen research directors from Germany's Max Planck Society met with European Commission officials for a frank exchange of views and information.

At one point, the years of frustration with the commission's successive, and increasingly Byzantine, Framework research programmes bubbled over, briefly pushing a mood of seething courtesy into undisguised rage. One of the scientists spat out his view that money is distributed in the current, sixth Framework programme (FP6) according to the rules of "lottery and lobbying". Achilles Mitsos, director general of the research commission, responded angrily to these L-words, barking his refusal to respond to "this kind of allegation" in the tone of an angry parent driven beyond reason.

Many researchers have been at loggerheads with the commission throughout successive Framework programmes, which have all too often been painful to join. Not that they lacked vision. FP4 came with a grand vision of integrating basic and industrial researchers from rich and poor countries. The messianic vision of FP5 was to plunge the newly integrated community into complex socio-economic problems. And the underlying vision of FP6 is the European Research Area, stimulated by an emphasis on building fewer but larger and well-funded mega-collaborations with such names as Networks of Excellence or Integrated Projects.

The current Framework programme is taking the application procedures to new extremes of complexity and anxiety. With each four-year programme shifting massively in philosophy, scientists have had to learn new rules of play. They must fill in ever thicker dossiers of forms, now so complex that many prefer to pay consultants to do the job, and twist their brain around near-surreal New-speak to justify their research — be it muscle physiology or new materials — in terms of gender relevance, for example.

But many researchers apply nevertheless, because their financial need is greater than their mistrust of the system, and this is why many also put up with success rates that have often been unacceptably low. Mitsos says that over-subscription to Framework programmes is a sign of success. He is mistaken. It is a sign of the need for European-level funding for basic research which does not exist elsewhere.

Central funding

For researchers, the commission serves, understandably but also unfairly, as the whipping boy. It is the executive arm of the world's most complex political grouping, the European Union (EU), and as such is bound by rules set by its wilful political masters, the Council of Ministers and the directly elected European Parliament. Many of the research commission's bureaucrats are themselves scientists who have bent over backwards to shoehorn basic research into programmes, which, they are instructed from above, must directly address the societal or economic needs of the EU citizen. But however they try, they will never be able to reconcile satisfactorily the need for no-strings basic research with the EU's political and legal requirement to serve the citizen.

So we can welcome the commission's initiative, announced by Mitsos in that small room after the fireworks and the wound-licking, to launch a serious political discussion on a realistic mechanism of funding basic research at the European level — one that researchers have long been calling for (see page 487). Importantly, the concept that he outlined foresees funding of such research by the commission, which would for the first time distribute money on the basis of scientific excellence alone, as judged by a peer-review system run by the scientific community.

The target, everyone knows, is the creation of a European Research Council (ERC), although that phrase will be introduced formally only when the concept has political backing. Some want to see an ERC fed by money from the budgets of national research councils, fearing that money given by the commission could never really be free from bureaucracy. But the commission seems to be sensitive to the issue and confident that it can be done. Moreover, it is in any case not clear that national research councils would be legally able to switch parts of their own budget to an international pool.

Viable solution

The commission's plan has a chance of being put into operation by 2007 if politicians play ball, in time for FP7. Its enactment is contingent on the commission receiving the major increase in research budget that is foreseen in the next round of financing. If the money and political will converge to launch a commission-funded ERC, will Europe's basic researchers still need the Framework programmes as well? Yes, because despite the general frustrations, many of the programmes selected for funding in FP6 do in fact fit well with the criteria, are of the highest quality and, like the vaccine-development programmes, are so large and wide-ranging that it is hard to imagine them being funded by a different mechanism.

To the mentors of the initial ERC concept — notably the heads of the Scandinavian, Dutch and German research councils — the idea that their bold vision might end up in the clutches of EU bureaucracy is no doubt a cause for dismay. They should be glad, however, to have the commission on their side, rather than working against them, which would nip progress in the bud.

The new European Constitution, if finalized as expected before the end of the year, will give the commission more power to try new approaches in research policy, without the full consent of the member states. Mitsos has made it clear that the commission is willing to grasp the opportunity. He has also clearly indicated that the winner will be basic research.

This does not necessarily mean that the commission has usurped the idea of an ERC, or that it intends to superimpose its bureaucratic structures on it. The chances are still good that an ERC will become a body that is legally and financially accountable to the commission but technically, scientifically and administratively autonomous. Europe's scientific leaders should be politically wise enough not to dismiss prematurely a compromise that could turn into a viable solution. ■





From the ruins
Iraqi academics lay plans for national science academy
p484



Gun ship
Outcry as US research vessel sails under flag of war-torn Liberia
p485



Deep trouble
Treaties powerless to avert severe climate change, say experts
p486



Sunny forecast
Swiss team aims for first solar-powered tour of the globe
p489

France and Japan lock horns in battle to host research reactor

Declan Butler, Paris

Japan and the European Union (EU) will this week contest rights to host the US\$5-billion ITER fusion experiment, after France saw off Spain's bid to be Europe's candidate. But uncertainty remains over whether the world's largest planned scientific project will proceed at all, researchers in several countries say.

Spanish negotiators had wanted the project's international partners to consider its own site near Barcelona, as well as France's site at Cadarache, near Marseille. But the EU last week unanimously backed the French site to compete with Japan's site at Rokkasho-Mura, 600 km north of Tokyo. A fourth site in Canada has effectively been withdrawn from the competition.

The ITER partners — the EU, Japan, Russia, Canada, China, South Korea and the United States — could choose the final winner as early as 5 December, when they meet in Vienna. Otherwise, a decision might be made when their representatives meet again in Washington on 16 December.

The ITER reactor aims to reproduce the thermonuclear fusion that powers the stars, heating a mix of deuterium and tritium, two isotopes of hydrogen, to 200 million °C, in a doughnut-shaped reactor called a tokamak. The plasma mix would be suspended by superconducting magnets, while a fusion reaction generates 410 MW of power for up to 500 seconds. ITER would be the first such reactor to sustain a power output comparable to the power that might be required in future fusion-based power stations.

Europe's choice of France could complicate the position of one partner — the United States — in what officials say have already been tense negotiations. The United States was officially 'neutral' on the European site decision, but Bush administration officials made little secret of the fact that they favoured Spain — an ally in the Iraq war — over France (see *Nature* 423, 211; 2003).

Scientists are keen that such political considerations stay out of the site choice, but as one US fusion researcher says: "US-French relations are strained. The issue is looming out there; you have to deal with it." Officials involved in negotiating the EU bid are



Prime minister Jean-Pierre Raffarin (right) and cabinet members examine France's plans for ITER.

equally keen to move on. As one of them puts it: "Cadarache is no longer a French candidate, it's a candidate of the EU, where there are strong US allies. So the United States wouldn't just be hurting one member state, but all."

But the negotiations have no simple voting rules, and officials and researchers say that the project probably needs a consensus from all partners to proceed. Failure in this month's talks could result in the entire project unravelling, they add. A breakdown in talks is "the lingering worry in the background", says one US fusion scientist. "I can see some of the rocks in the water."

A deal will involve consensus on engineering, costs, culture and politics, though perhaps not in that order. On paper, the French site has some advantages, such as existing on-site industrial and scientific nuclear capacity. But several fusion pundits said that Japan has the edge over Europe, because they think that it may be willing to spend more cash and it is close to ITER's new partners Korea and China. The amount of money each party is prepared to put on the table this week could sway the choice, they say.

Some US scientists privately told *Nature* that they would prefer to spend the next ten

years working on the project in the south of France than in Japan. But any such individual preferences will probably be outweighed by broader, national interests, says Gerald Navratil, a fusion scientist at Columbia University in New York. "Japan may be the consensus site in that all are Pacific-rim nations except the EU," he says. "If political considerations include effective involvement of Korea and China and strengthening US ties in Asia, the United States may well go this way."

"We would hope to have their support because of our proximity," says Hiroshi Kishimoto, the ITER spokesman for the Japan Atomic Energy Research Institute.

For their part, scientists seem less concerned with the choice of site than with the risk that any disagreement could scupper ITER itself. "The highest priority for the US scientific community is that a decision should be made to move forward," says Robert Goldston, director of the Princeton Plasma Physics Laboratory in New Jersey, adding that he feels that the partners can build it whichever site is chosen. "Many of us in the community have worked for years, some for decades to bring us to this point," adds Navratil. "We may finally pull this off."

C. PARIS/AP

Shortcomings halt study of Swiss cancer vaccine

Quirin Schiermeier, Munich

A Swiss clinical trial of a vaccine against skin cancer has been scrapped after a government-backed investigation found "serious shortcomings" in its conduct.

The inquiry, which released its results on 28 November, says that the study's methods and medical examinations were insufficiently documented, its protocols inadequate, and that several patients were treated without prior written consent.

But Frank Nestle, the dermatologist at the University of Zurich who led the study, will not face sanctions because the Swiss federal health office's inquiry found that no scientific data had been manipulated and that patients had not been put at risk.

In 1998, Nestle reported tumour regressions in 5 out of 16 melanoma patients vaccinated with a fusion of dendritic cells from the immune system and tumour cells (F. O. Nestle *et al.* *Nature Med.* 4, 328–332; 1998). This success rate dropped substantially after more patients were included in the study.

But, according to the inquiry, only the earlier, more promising results were displayed on the clinic's website. This was the case until February, when the University of Zurich responded to complaints and set up an internal investigation of the study. The clinical trial was suspended on 17 April, and the Swiss federal health office commissioned its inquiry.

The Internet material was meant primarily for physicians, Nestle says. But he admits that he cannot exclude the possibility that new patients — including some non-Swiss nationals who were charged an average of SFr9,000 (US\$7,000) to take part in the study — could have been persuaded to participate because of the information on display.

Nestle says that he did not intend the website to be an advert for the study. "As an important cancer centre, we were allocated more than enough patients," he says. In addition, he notes that treatment costs were administered by the university's hospital management, which did not break existing guidelines. "There was, of course, no personal gain," he adds.

The initial Zurich work led to a larger, multicentre study of dendritic cells as a treatment for skin cancer, coordinated by the German Cancer Research Center in Heidelberg. Despite doubts about cancer vaccines now raised by the Zurich study, trials in Germany and Switzerland on more than 100 patients will continue. ■



Royal Society officials meet with Hussain Al-Shahristani (right) to discuss next steps for Iraqi science.

Iraqis draw up blueprint for revitalized science academy

Jim Giles, London

A dozen Iraqi scientists and engineers met with officials of the Royal Society in London last week and agreed to set up an Iraqi academy of sciences. They said that the academy would help to develop a research strategy for the country.

Although most of the Iraqi participants work in Iraq, the meeting was held in London for security reasons, organizers said. Hussain Al-Shahristani, an exiled chemical engineer who initiated the meeting, says that the organization will be open to established scientists and engineers, and will be independent of the government. He is hoping to attract "a few hundred thousand pounds" from Western sources to set up the body, whose first full meeting is planned for next November in Baghdad.

A former senior adviser to the Iraqi Atomic Energy Commission, Al-Shahristani says that he was imprisoned and tortured during Saddam Hussein's regime for refusing to work on an atomic-weapons programme. He escaped from Iraq in 1991, when Baghdad was bombed during the first Gulf War, and is currently a visiting professor at the University of Surrey, UK.

Al-Shahristani suggests that establishing environmental studies of biological, chemical and radiological pollution should be a priority for the academy. In particular, he would like to find out whether the pollution, some of which was caused by weapons-testing programmes, is linked to an increase in cancer

rates that has been reported in the country.

Other exiled participants said that they were unsure if they would ever return to Iraq. Farhan Bakir, former personal physician to Saddam Hussein, left in 1981 after being forced into early retirement. Bakir says that he feared for his safety.

"People were being killed and arrested," he recalls. "I was insecure, but so was everyone in Iraq." Bakir, currently at the Western University of Health Sciences in Pomona, California, says that he would like to go back, but will wait until things are more settled. "I'm over 70," he jokes. "I'll take an advisory position at this age."

The Iraqi Academy of Sciences that operated under Saddam had few international contacts, and has now ceased to function. Critics say that it was directly controlled by officials close to Saddam. Huda Salih Mahdi Ammash, an Iraqi microbiologist who was reportedly a senior figure in the country's biological-weapons programme and is now in US custody, ran the academy from the mid-1990s.

But nuclear, chemical and biological weapons pose a dilemma for the new academy. Scientists at the London meeting said that those involved in developing weapons of mass destruction would not be allowed to join. But Al-Shahristani acknowledges that some scientists had little choice about the area they worked in. "People know who was forced into weapons work," he says. "We'll look at it on a case-by-case basis." ■

Ship row flags up funding of war in Africa

Rex Dalton, San Diego

The flag that flutters atop the *JOIDES Resolution*, one of the world's premier scientific ships, does not represent any of the more than 20 nations that have funded her decades of journeys on the seven seas. Instead it is the banner of Liberia — a West African country that has diverted funds from its fat ship-registration coffers to fund arms purchases for civil wars.

The *Resolution*, like many other ships, flies under this flag of convenience in order to limit taxes, while operating under fewer regulations for safety and crew working conditions.

But human-rights organizations have discovered in recent years that Liberia and neighbouring countries have used money from shipping registration to fund conflicts noted for their use of child soldiers. This has prompted some commercial ship owners to shift their registration elsewhere. Now a growing number of scientists are hoping that the owners of the *Resolution* will follow suit.

The *Resolution's* scientific cruises are managed by the Joint Oceanographic Institutions (JOI), a Washington-based consortium of 18 public institutes that receive funds from the US National Science Foundation (NSF). In September, the JOI won a ten-year, \$625-million contract from the NSF to manage the US component of the Integrated Ocean Drilling Program (IODP), an international drilling endeavour (see page 492).



Flag of inconvenience? Cash spent to register the *Resolution* in Liberia ends up funding wars in Africa, critics say.

The JOI learned of the concerns about Liberian registry at least two years ago, when a former third mate from the *Resolution* wrote to US Senator Michael Enzi and Representative Barbara Cubin, both Republicans from Wyoming, complaining about the use of NSF funds for a Liberian ship registry. The mate branded the practice un-American, and claimed that it exploited crewmen from developing nations. "The NSF and the American taxpayer should never fund a ship that does not fly the flag of the USA," the mate wrote.

JOI president Steven Bohlen says that the seaman was a "disgruntled employee" who was discharged because of limited skills. And both the JOI and the *Resolution's* owners

insisted that the ship was operated safely under Liberian rules. The members of Congress took no action. Neither did the JOI or the NSF, whose officials say that the agency does not have a position on the issue.

But as the *Resolution* prepares to sail for the IODP, concerns are re-emerging from scientists, who wish to remain anonymous, that scientific funds from the United States, Japan and Europe are continuing to go to

Liberia. They say that few researchers associated with the IODP are familiar with the issue, but that many of those who are feel strongly about a change in registration.

Bohlen says that his agency has been too busy developing the IODP to address the issue of ship registry. But Guy Cantwell, a spokesman for Transocean, the world's largest offshore drilling firm, which is based in Houston, Texas, and co-owns the *Resolution*, says: "We are aware that this is an issue in the scientific community." After *Nature* contacted Transocean, Cantwell said that the firm is considering switching the *Resolution's* registration to a more politically palatable low-tax nation, such as the Marshall Islands in the Pacific. ■

Budget cuts force Hong Kong to reduce salaries

Carina Dennis, Singapore

Scientists in Hong Kong are bracing themselves for a wave of salary cuts as the government reduces university funding next year by nearly 10% from last year's figure.

Staff salaries are expected to be hit hardest. Research funding, which is overseen by the Research Grants Council, has so far been spared cuts. But researchers worry that low salaries will drive some people away, leaving fewer to teach and giving the remaining staff less time for research.

On 26 November, the University Grants Committee issued a statement outlining the 9.7% reduction for 2004–05. The cuts, which total about HK\$1.1 billion (US\$140 million), are part of a strategy of the Hong Kong Special Administrative Region government to reduce expenditure to soothe the region's troubled economy, which was battered by the SARS epidemic.

University funding in Hong Kong is usually decided for three years at a time — the last budget covered 2001–04. But the government decided to set the upcoming budget for only a single year, following a review of the higher-education sector that recommended that the region's eight universities diversify into their areas of strength.

"We decided to give the University Grants Committee a year to prepare new funding models and to give the universities time to focus on differentiating their roles," says Arthur Li, Hong Kong's secretary for education and manpower. In the meantime, the government is encouraging universities to make up some of the shortfall through fundraising, with a HK\$1-billion scheme to match funds obtained from other sources.

The government will not comment on whether there will be additional cuts in

the next three-year budget, but many researchers are concerned about the future. "A lot of young people don't know if their contracts will be renewed or what their career future will be in Hong Kong," says Kathy Cheah, who heads the University of Hong Kong's biochemistry department.

But not everyone is being driven away by the budget problems. Lap-Chee Tsui took up the position of vice-chancellor at the University of Hong Kong last year after serving as chief geneticist at the Hospital for Sick Children in Toronto, Canada. "The main reason I came back was to contribute to the education sector. I didn't expect the budget situation to be as it is but I'm happy to be part of solving the problem," he says.

"Even though it is a difficult economic time, I hope that the government won't be so short-sighted as to cut our roots so we can't grow in the future," adds Tsui. ■

Scientists go to jail to crack substance abuse

Emily Singer, Washington

The US National Institute on Drug Abuse (NIDA) is embarking on a programme to study the treatment of drug abuse among prisoners in the United States — and to educate judges about the neurobiology of addiction.

NIDA officials hope that the programme will help to break the cycle of drug abuse, crime and imprisonment by finding better ways of selecting prisoners for different drug treatment options. The institute plans to spend \$30 million on the programme over the next five years, conducting studies at ten prison sites across the country.

At a panel discussion at the Society for Neuroscience meeting in New Orleans last month, NIDA director Nora Volkow referred to the programme, which she dubbed “NIDA goes to jail”, as an example of how scientists can interact with societal issues.

Drug abuse is a vast problem for the US criminal-justice system: 60% of inmates have been convicted of drug-related offences. Nine out of ten parolees with a history of drug abuse return to drugs, according to Douglas Marlowe, a psychologist and criminal-justice researcher at the University of



Behind bars: three out every five US prisoners were convicted of drug-related offences.

Pennsylvania in Philadelphia, and many are re-arrested on drug-related charges.

The programme is still at the planning stage, NIDA officials say, but is likely to begin with a survey of prison administrators, clinicians, wardens and parole officers at the ten sites to establish existing treatment patterns.

NIDA will also sponsor clinical trials, beginning next year, of different ways of assessing prisoners for treatment before their release. By measuring factors such

as motivation for treatment and levels of hostility, anxiety and risk-taking behaviour, researchers hope to match patients with the best type of treatment, says Wilson Compton, a psychiatric epidemiologist and head of prevention research at NIDA. “People haven’t focused adequately on this population,” he says.

The institute also plans to step up its efforts to educate judges and police officers on the biology of addiction. Officials at Treatment Accountability for Safer Communities (National TASC), an organization based in Alexandria, Virginia, that helps to treat drug addicts, say that NIDA has been consulting with it on how best to do this. A National TASC report on current education efforts should be ready within a month, and will help NIDA to plan a pilot education programme to be conducted in North Carolina.

Marlowe thinks that this type of education is badly needed. “Judges have heard that drug addiction is a disease, but they don’t know what that means,” says Marlowe. “They don’t understand why a person relapses after being drug-free in jail for six months. If they know how cravings are triggered by environmental cues, they would understand.” ■

Climate study highlights inadequacy of emissions cuts

Quirin Schiermeier, Munich

As representatives from 188 countries gather in Milan, Italy, to work through the latest round of negotiations for the Kyoto Protocol on limiting greenhouse-gas emissions, German researchers have released a study showing that “intolerable” levels of climate change are more likely than decision-makers may realize.

In a report published on 25 November, the German Advisory Council on Global Change contends that the world can tolerate a rise of up to 2 °C over pre-industrial levels. Beyond this, the effects of climate change on society would become too severe, they say. This would be mainly due to sudden phenomena such as the possible irreversible disintegration of large ice sheets, or abrupt disturbances to the North Atlantic Ocean’s currents and to monsoons in Asia.

Global mean temperatures have already increased by 0.6 °C since 1900. By the end of the twenty-first century, temperatures may increase by a further 1.4–5.8 °C, according to the latest projection of the Intergovernmental Panel on Climate Change.

Even the lower limit of this estimate will pose an intolerable threat to human health, food and water supplies, economic

development and natural ecosystems in many parts of the world, says the German report.

But it seems unlikely that political action will be able to keep temperature increases down. The 1997 Kyoto Protocol has yet to be ratified, as Russia has not decided whether to sign it. But even if the protocol comes into effect, the cuts that it prescribes will be insufficient to hold climate change to a ‘tolerable’ level, the report says.



Flood warning: global warming is widely expected to disrupt the monsoon in Pakistan.

It concludes that global carbon dioxide emissions would need to be curbed by 45–60% by 2050 compared with 1990 levels to avoid dangerous climate change. The Kyoto Protocol calls for average cuts of 5% in industrialized countries by 2012, and even that may be unrealistic.

The report will be presented in Milan at the ninth Conference of the Parties to the United Nations Framework Convention on Climate Change on 10 December. But it is unlikely to have an immediate impact on Kyoto negotiations, say the report’s authors.

“These findings reinforce our position,” says Karsten Sach, head of the German environment ministry’s international climate-protection department, and leader of the German delegation in Milan. Germany is committed to a reduction of 40% in its emissions by 2020, he says, regardless of whether the Kyoto Protocol comes into effect. Other European countries, including Britain, France, Sweden and the Netherlands, have similarly set their own goals.

Plans for a climate-protection strategy beyond Kyoto are not yet on the table, but some analysts expect the Italian host delegation in Milan to outline possible next steps. ■

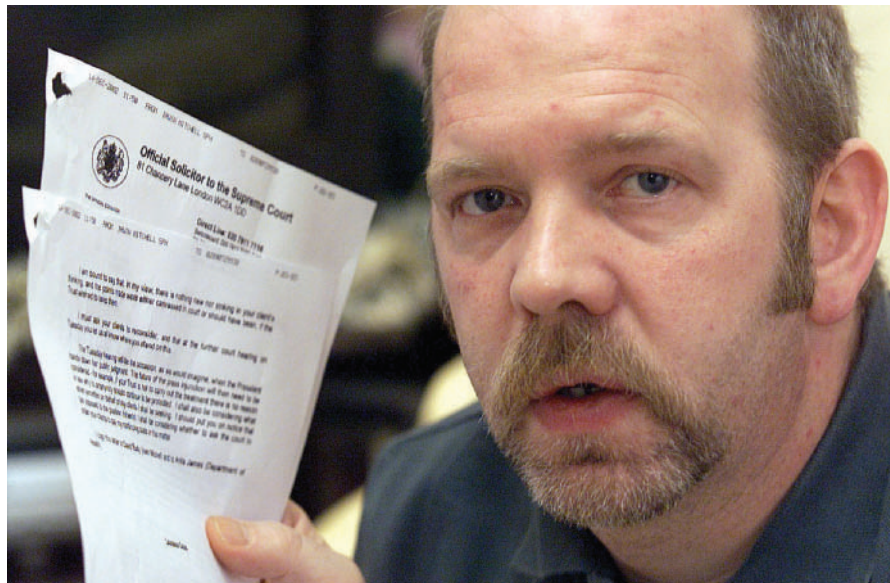
Rapid drug trial offers hope to CJD patients

Jim Giles, London

Britain's Medical Research Council is set to embark on a rigorous clinical trial of a treatment for Creutzfeldt-Jakob disease (CJD). The trial will initially assess the effects of only one drug — a compound that few experts expect to be effective. But the trial has been designed to allow newly discovered drugs to be rapidly tested, a move that doctors hope will sort out the chaos that has been caused in the past by dying patients' demands for access to untested treatments.

The study, which is expected to receive final ethical approval within the next few months, will look at treatments for all forms of CJD. This includes variant CJD — thought to be contracted through eating meat infected with bovine spongiform encephalopathy (BSE) — and sporadic CJD, a rare condition with no known cause. About 900 people in Britain have died from these diseases since 1990.

Patients have been eager to get their hands on any possible treatment for these conditions. That includes quinacrine — the drug that will be tested first in the trial. Quinacrine, which was initially developed as a malaria treatment, was shown in lab tests with cell cultures to inhibit the formation of prions — the proteins thought to cause CJD (C. Korth, B. C. H. May, F. E. Cohen and



Court battle: Don Simms of Belfast, UK, with the papers allowing his son Jonathan to start pentosan treatment last January. Jonathan has improved slightly, but the drug is not included in the new trial.

S. B. Prusiner *Proc. Natl Acad. Sci. USA* **98**, 9836–9841; 2001). Stanley Prusiner, who led that study and also won a 1997 Nobel Prize for his work on prions, backed the use of quinacrine to treat CJD patients and, at around the same time, reported that the symptoms of a British CJD sufferer had improved after taking the drug.

Media coverage, as well as Prusiner's endorsement, led to a scramble to use the drug before any proper trial had been done, regardless of the fact that useful effects in cell cultures are often not replicated in animal tests. "It was premature to state that it was good for patients," says Adriano Aguzzi, a neuropathologist at University Hospital Zurich in Switzerland.

Since then, hopes for quinacrine have dimmed. Evidence from human CJD sufferers suggests that it has little effect, says Richard Knight, a neurologist at the National CJD Surveillance Unit in Edinburgh. Animal experiments carried out since then have also failed to show unambiguously that the drug works (P. Brown *Neurology* **58**, 1720–1725; 2002).

Nevertheless, Knight says that the British trial will be welcomed by patients, as it should reveal whether quinacrine has any effect, however minor. More importantly, he adds, the trial will allow rigorous assessment of other drugs as soon as any potential benefit is revealed.

The trial will not put an end to the use of untested drugs, however. At present, one drug in demand from CJD patients is pentosan polysulphate. Unpublished work carried out by Katsumi Doh-ura, a prion researcher at Tohoku University in Sendai, Japan, suggests that this compound can delay the onset of symptoms in mice infected with disease prions.

Patients have taken the battle to court and won the right to be given the drug. But because it has to be injected into the brain, and is linked to side-effects such as impaired blood clotting, it has been considered too dangerous to include in the British trial.

Europe eyes merit-based agency

Alison Abbott, Brussels

The European Union (EU) last week edged closer to setting up a new research agency that would be run by scientists and would give out grants purely on the basis of scientific merit.

Achilleas Mitsos, director-general of the European Commission's research directorate, said last week that the commission would publish a position paper within the next few weeks for a "major programme in basic research where excellence is the sole and exclusive priority, and where evaluation will be in the hands of the scientific community".

Until now, the EU has funded only small amounts of basic research under its Framework research programmes, which are designed primarily to support applied research. Most basic research is supported by national research councils.

Mitsos was speaking at a meeting last month in Brussels of research directors from Germany's Max Planck Society, members of the research commission and research commissioner Philippe Busquin.

The concept that Mitsos proposed is close to the advice given by an expert group, which recently called for a European Research Council to be set up, financed with EU money but run independently of European Commission officials (see *Nature* **425**, 440; 2003).

Mitsos said that the commission will not use the term 'European Research Council' in its document, pointing out that although the concept has been intensively discussed in the scientific community during the past couple of years, it has not been discussed in political circles.

The commission's paper will, however, be discussed in spring by the European Council of Ministers and by members of the European Parliament.

"We have to do this in a clever way," Mitsos said. "The commission cannot afford to propose something that would be rejected — we have to prepare the ground." He hopes that the paper will persuade the council to ask the commission to prepare a formal proposal for a basic research council in the second half of next year.

United States and India forge military research deal

New Delhi India and the United States are set to sign an agreement that will allow their defence scientists to share classified military research.

Existing arrangements allow the two nations to carry out joint military exercises and arms purchases. The updated agreement will let defence labs exchange scientists and information. Indian scientists say that they are keen to learn more about electronic warfare and anti-ballistic-missile defence, and add that their US counterparts

may benefit from India's experience of warfare in extreme environments. India has expertise in mountain warfare and has developed herbal drugs for treating frostbite and pulmonary oedema, a condition caused by lack of oxygen. It also has treatments that boost stamina and that provide protection against chemical and biological attacks. Indian psychologists have also studied the use of yoga and meditation to combat stress.

The draft agreement, which has been taking shape since February, was finalized at a meeting in New Delhi last month between India's Defence Research and Development Organisation and a 15-member delegation of US defence scientists. It now awaits signatures from top officials on both sides.

Organic brewery ferments opposition to GM crops

San Diego An organic brewery in California has successfully petitioned to get a question about banning the growth of genetically modified crops included in a forthcoming county ballot.

The campaign is based in rural Mendocino County, about 200 kilometres north of San Francisco. Els Cooperrider, a campaign leader and co-owner of the Ukiah Brewing Company, says that he has had numerous requests about how to duplicate

the initiative, both in other US counties and in countries such as Brazil.

No transgenic food is currently produced in Mendocino County, a region known for its liberal political views, organic farming and illicit marijuana plantations.

The county's 45,000 voters are expected to vote on 2 March on whether to ban the growing of transgenic crops. Campaigners say that representatives of the biotechnology industry are considering sending staff to Mendocino to lobby against such a ban.

Europe launches attack on bacterial resistance

London The European Union (EU) has announced plans to spend E13 million (US\$16 million) on studies of antibacterial resistance. The decision was made as investment in antibiotics research wanes (see *Nature* 425, 225; 2003) and the incidence of multidrug resistance in bacteria is increasing.

Details of the two planned studies were released on 28 November at the European Conference on Antimicrobial Resistance in Rome. One study will focus on how resistance to a major class of antibiotics, known as β -lactam antibiotics, arises when infections spread within hospitals and communities.



Future aim: Indian troops will soon be sharing survival techniques with US defence experts.

R. A. VEENDRAN/AFP/GETTY IMAGES

Streptococcus pneumoniae, which causes pneumonia and meningitis, is the subject of the second study. Researchers will examine how the bacterium can grow and spread in the presence of some antibiotics. Both projects are funded by the European Union's Sixth Framework Programme for research.

Balloon adventurer seeks to broaden horizons

Munich Engineers at the Swiss Federal Institute of Technology in Lausanne are due to begin work on the Solar Impulse — a solar-powered aircraft that they hope will complete a non-stop flight around the world.

The team was asked to look into the project by Swiss adventurer and psychiatrist Bertrand Piccard, who, with the help of the same researchers, became the first person to successfully circumnavigate the globe in a hot-air balloon in 1999.

The institute has conducted a feasibility study and has agreed to take on the research. Michel Declercq, the institute's dean of engineering, says that the biggest challenge will be to keep the plane in the air overnight. One solution is for it to climb to 10,000 m during the day, giving it enough altitude to descend slowly during the night until the solar cells kick in again at dawn. To achieve this, the researchers will have to



Blue-sky thinking: an artist's vision of the Solar Impulse. The real version could take off in 2007.

design devices for efficiently capturing and storing solar energy. They expect the first manned flights on a prototype to take place by 2007.

US names first homeland security centre of excellence

San Francisco Details of the first US centre of excellence in homeland security were unveiled last week.

The University of Southern California in Los Angeles will host the institute, known as the Homeland Security Center for Risk and Economic Analysis of Terrorism Events. The centre will assess the risk of attacks on critical infrastructure such as electrical and telecommunications networks, and develop

strategies for responding to such attacks. The project will receive \$12 million over three years.

The centre is one of ten planned by the Department of Homeland Security. Details of the next centre, which will focus on the threat of agricultural bioterrorism (see *Nature* **421**, 106–108; 2003), will be released later this month.

Earth maps put environment treaties into perspective

Paris A map showing how Earth's surface changed during the 1990s was unveiled at the Global Monitoring for the Environment and Security meeting in Baveno, Italy, last week. The data will be used by the United Nations in its efforts to track the impacts of environmental conventions over the past decade, and by institutions such as Météo France to help with weather forecasts.

The Global Land Cover 2000 map shows the extent of 22 types of land cover — from farms and cities to deserts, tropical forests, wetlands and snow fields — and is based on data collected between November 1999 and December 2000 using the French SPOT 4 satellite. It shows many changes since the last map was compiled using data from 1992 and 1993, including the disappearance of some six million hectares of tropical forest.

Take a cell, any cell...

Can an adult human cell be turned back to an embryonic state without the need for cloning? If so, ethical objections to personalized regenerative medicine would be swept away. Carina Dennis reports.

It's the modern equivalent of alchemy: turning any nondescript human cell into biomedical gold. The treasure in this case is a cell that will resemble an embryonic stem (ES) cell, which can grow into any type of tissue. This cellular alchemy might one day provide the means to repair a failing body with grafts that are derived from the patient's own cells — and so won't be rejected by their immune system.

So far, there's only one proven method of turning back a cell's developmental clock: take an adult cell and fuse it with an unfertilized egg cell stripped of its own genetic material to create a cloned embryo. After growing the embryo for a few days in a culture dish, you can harvest its ES cells.

Huge technical and logistical challenges face those trying to turn this 'therapeutic cloning' into clinical practice. And even if these hurdles can be overcome, many people feel moral repulsion at the idea of creating a human embryo for the sole purpose of using its cells in a medical procedure. In some countries, such as France and Germany, the practice is already banned.

But if adult human cells could be reprogrammed into becoming ES cells, without creating an embryo, these objections would

wither away. In any case, donated human eggs are in very short supply, and are in demand for *in vitro* fertilization. "We need to find an alternative," says Paul Verma of the Monash Institute of Reproduction and Development in Melbourne, Australia.

The unknown factor

With this in mind, some researchers are experimenting with other cell types that seem to share eggs' capacity for developmental reprogramming. In the long run, however, the best solution will be to identify the biochemical factors sloshing around in an egg's voluminous cellular soup that can wipe an adult cell's slate clean and start afresh.

One place to look for these factors is in ES cells themselves. Fusing adult mammalian cells with ES cells, or with the embryonic cells that ultimately give rise to sperm and eggs, can make adult cells take on some of the characteristics of ES cells^{1,2}. "Experiments suggest that ES cells and eggs have similar factors that can reprogramme adult cells," says Azim Surani of the Wellcome Trust/Cancer Research UK Institute of Cancer and Developmental Biology in Cambridge, whose team reported early work of this kind.

The problem is that ES cells are small and difficult to work with, and cannot readily be stripped of their genetic material in the same way that eggs can. So, reprogramming experiments involving ES cells typically end up with hybrids containing chromosomes from the ES cell as well as from the cell to be reprogrammed. This defeats the eventual purpose, which is to create stem cells that are an exact genetic match to the patient. It also raises fears that such genetically abnormal cells might become cancerous. "The next challenge is to eliminate the ES-cell genome," says Takashi Tada of the Institute for Frontier Medical Sciences at Kyoto University in Japan, a former member of Surani's team.

Verma and his Monash colleagues may have the

answer. In work still in progress, they have fused pairs of ES cells to create large cells with nuclei that are double the normal volume. They then fuse one of these doubled-up cells with the cell to be reprogrammed, but prevent the nuclei from joining together by adding a chemical inhibitor that disrupts the hybrid cell's cytoskeleton. The heavier ES-cell nucleus can then be spun out of the hybrid cell using a centrifuge. Whether the remaining nucleus is effectively reprogrammed remains unclear, but Verma is optimistic. "It's early days, but we see gene expression changes that suggest rapid reprogramming in the hybrids," he says.

At the University of Oslo in Norway, Philippe Collas and his colleagues are trying to get round the problem by using ES-cell extracts, rather than the cells themselves. They punch holes in the outer membrane of adult cells using a bacterial toxin, then douse them with the cellular juice squeezed from embryonic cells and, after 30 minutes, allow the holes to reseal.

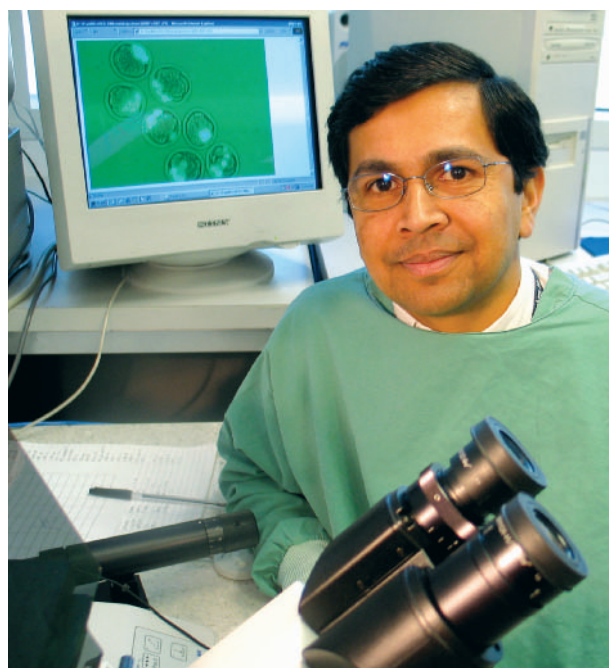
Collas's team originally used this approach to confer characteristics of immune T cells on fibroblasts — a common cell type in skin and other connective tissues³. The researchers are now treating mouse fibroblasts with ES-cell extracts. "We are seeing some ES-cell-specific properties," Collas claims.

Not everyone is convinced. "I'm sceptical," says Rudolf Jaenisch of the Whitehead Institute for Biomedical Research in Cambridge, Massachusetts. If an adult cell infused with an ES-cell extract subsequently shows some molecular characteristics of an ES cell, he says, it may simply reflect the material it has taken up — it doesn't necessarily mean that the adult cell's genes have been reprogrammed to an embryonic state.

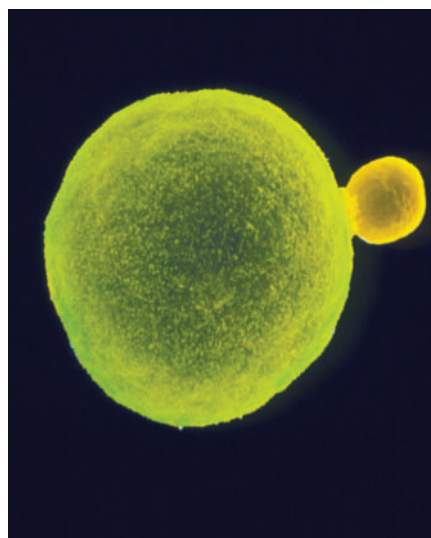
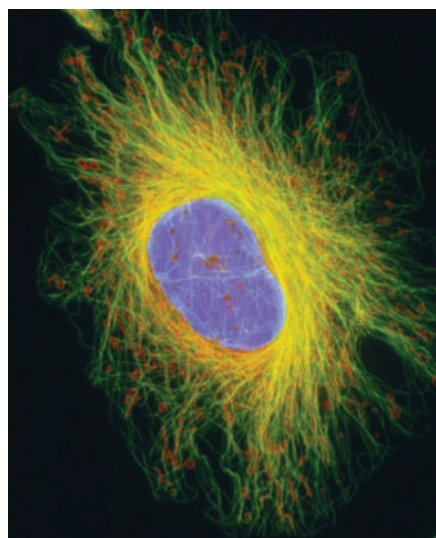
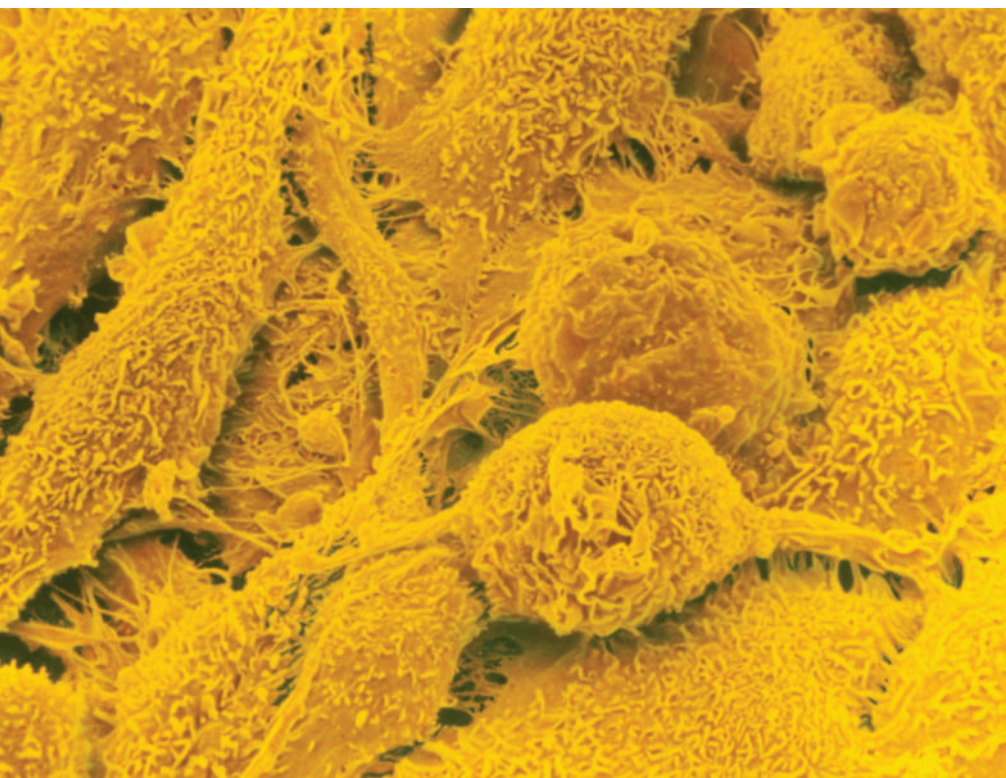
Back to the future

Another problem is that any solution that relies on ES cells to achieve developmental reprogramming does not remove the fundamental moral objection to therapeutic cloning, as the ES cells must originally come from a human embryo. It would be far more desirable to be able to turn back a cell's developmental clock by adding a defined cocktail of biochemical factors.

Several research teams are looking hard at the biochemical soup inside eggs and embryonic cells in the hope of pinning down these factors. The exact nature and number of the



Fusion biology: Paul Verma is looking for more efficient ways of persuading embryonic stem cells to reprogramme adult cell nuclei.



Going for gold: currently, deriving embryonic stem cells (top) from an adult mammalian cell, such as this fibroblast (above left), requires a cloning step. But a handful of research groups are now trying to achieve the same trick using reprogramming factors isolated from egg cells (above right).

factors is unknown. “But we can make some educated guesses,” says Surani. The chief suspects are molecules that have been shown to erase the previous memory of a cell’s state and to reconfigure the genome to reactivate long-silenced genes.

A cell arranges its genome like some of us organize our wardrobe. Active genes, like this season’s fashion, are kept handy and are readily accessed by the cell’s biochemical machinery. But genes that are rarely used are tucked away in the genome’s attic like last winter’s overcoat.

DNA inside each cell is wrapped snugly around histone proteins and packed up into

a structure called chromatin. In regions of the chromosomes where genes are active, the chromatin is loosely packed so that genes can be accessed; dormant genes tend to be buried in dense chromatin. Reprogramming a cell’s nucleus involves dusting off mothballed genes by unwinding chromatin and modifying histones. Methyl groups, which can alter a gene’s activity, are also added or removed from the DNA during reprogramming.

Surani is looking for changes in histone modification and DNA methylation in adult cells that have been reprogrammed by fusing with embryonic cells. Although he says it is too early to name names, “we expect to see

patterns of modification that suggest specific candidates”.

Other researchers are looking for reprogramming factors in egg cells — a search spurred on by experiments conducted by John Gurdon, also at the Wellcome Trust/Cancer Research UK institute, which have shown that human and mouse nuclei can be reprogrammed when injected into frog eggs⁴. “Reprogramming factors found in frogs are likely to be conserved in other species, including humans,” says Gurdon.

Inside the nucleus

Developmental biologists have carried out cloning experiments with frog eggs for decades. They are huge, compared to mammalian eggs, and abundant. “You can get a litre of eggs from 20 frogs every day,” says Nobuaki Kikyo of the Stem Cell Institute of the University of Minnesota in Minneapolis. Kikyo has already fished out several factors from frog eggs that may be involved in reprogramming, including a complex of proteins that is involved in loosening up chromatin⁵.

Most recently, Kikyo’s team discovered a factor that induces dramatic structural changes in the nucleus during reprogramming. Inside the nucleus is a tiny compartment called the nucleolus, which contains a reservoir of enzymes that repair and maintain the genome. During reprogramming, the nucleolus breaks down and then reassembles. By trawling through frog-egg extracts, Kikyo’s team has pinned down a complex of proteins that causes the nucleolus to disassemble⁶. “We imagine that some factors in the nucleolus need to be released in order for the nucleus to be reprogrammed,” he says.

The key to identifying reprogramming factors is knowing how to look for them. “It’s important to have a good assay system,” says Gurdon, who is using the human gene *Oct-4*, which is switched on in embryonic cells, as a ‘bait’ to fish out reprogramming factors. If it becomes active when treated with a particular extract, that’s a strong sign that the extract contains a reprogramming factor.

Researchers hope that the number of essential reprogramming factors will be manageable small — otherwise it may never be possible to create personalized ES cells without resorting to therapeutic cloning. But if the cellular alchemists can find a handful of factors that does the trick, we all stand to be enriched by a biomedical gold rush. ■

Carina Dennis is Nature’s Australasian correspondent.

1. Tada, M., Tada, T., Lefebvre, L., Barton, S. C. & Surani, M. A. *EMBO J.* 16, 6510–6520 (1997).
2. Tada, M., Takahama, Y., Abe, K., Nakatsuji, N. & Tada T. *et al. Curr. Biol.* 11, 1553–1558 (2001).
3. Häkkelien, A.-M., Landsverk, H. B., Robl, J. M., Skälhög, B. S. & Collas, P. *Nature Biotechnol.* 20, 460–466 (2002).
4. Byrne, J. A., Simonsson, S., Western, P. S. & Gurdon, J. B. *Curr. Biol.* 13, 1206–1213 (2003).
5. Kikyo, N., Wade, P. A., Guschin, D., Ge, H. & Wolffe, A. *Science* 289, 2360–2362 (2000).
6. Gonda, K. *et al. Nature Cell Biol.* 5, 205–210 (2003).

YORGOS NIKAS/SP

GOPAL MURTHI/SP; P. MOTT/SP

Digging in

A new ship and a wave of funding will let scientists drill where they have never been able to drill before, from near the North Pole to the rocks lying beneath Earth's crust. Rex Dalton and David Cyranoski report.

Late next summer, when the Arctic pack ice has receded, three icebreakers will head north from Siberia to escort a drilling ship into the treacherous Arctic Ocean. Cutting through ice floes, the convoy will launch one of the most ambitious ocean-drilling projects ever undertaken, in an attempt to find a missing piece of the puzzle of our planet's climatic past.

Tentatively approved last month, the expedition will be one of the first projects of the new Integrated Ocean Drilling Program (IODP) — a scheme that will unite equipment and expertise from the United States, Japan and a consortium of 15 European nations. The programme aims to drill where no ship has drilled before, from the icy North Pole to deep inside a fault line off the coast of Japan, and to study the life forms they find deep beneath the sea floor along the way.

After a decade of planning, the IODP was officially born on 1 October. It replaces the Ocean Drilling Program (ODP), an enterprise of 20 nations that was phased out in September after nearly 20 years of expeditions. The ODP had only a single ship — the US vessel *JOIDES Resolution* — which sailed the globe, sinking holes into the ocean bottom and retrieving a huge volume of sediment cores (see 'Core values', overleaf).

Over two decades, the *Resolution's* bounty has helped to prove that asteroids such as the one that slammed into Mexico 65 million years ago — thought by some researchers to have caused the dinosaurs' demise — can create clouds of dust that settle all around the world. Sediment cores pulled from the Indian coast helped to put a date on when the Himalayas were born, as they revealed a mass of rock that was weathered by rain as the mountains were pushed up. And researchers have gained clues about Earth's ancient climate — the fossils found in layers of mud laid down on the ocean floor over millennia have revealed the temperature of those waters long ago.

But the *Resolution* couldn't go everywhere, and never mounted a campaign on the North Pole. In 1996, geologist Jan Backman, of Stockholm University in Sweden, and his colleagues took a lone icebreaker — operated independently of the ODP — and tried to drill a deep core on the Lomonosov Ridge, an undersea mountain ridge that meanders across the top of the world. The wind blew ice into the side of the ship, making it difficult even to stay still, and the team only managed to get a few, small

cores by dropping a weighted pipe off the back of the ship. "The ice does terrible things up there," says Backman. "When I came out, I said I would never do that again without the proper vessels."

Pole position

Now Backman has a second chance, as head of a mission that will return to Lomonosov in August. With US\$12 million from the IODP, Backman will this time have at his disposal satellite-borne radar that can see through clouds, real-time weather updates, three ice-breaking escort vessels, and an oil-industry drilling ship, on which the party will steam to a location 130 kilometres south of the North Pole. There Backman hopes to sink three holes into the ridge and retrieve 500 metres of cored sediments, which will give a glimpse of the past 50 million years. The sediments will be examined for minerals that have been transported by ice from nearby land, or fossils of tiny creatures that thrive in specific climates — tell-tale signs of when the pole



Illustrious career: the research drilling vessel *JOIDES Resolution* has sailed the world for more than two decades.

was free of ice, or completely iced over.

If the mission succeeds, it will be a coup for climatologists, for whom the North Pole's ancient climate is still a mystery. But the IODP has other ambitious projects in mind too. Along with its ice-breaking facilities, provided mainly by the European project members, it will also have a Japanese ship designed to drill deep into the Earth — the US\$500-million *Chikyu*. Unlike the *Resolution*, which was an old oil-exploration vessel converted for use by scientists, the *Chikyu*

has been designed from the keel up specifically for research. The ship has equipment for dating and analysing cores that is more comprehensive than that in most land-based labs, says Millard Coffin, a marine seismologist at the University of Tokyo's Ocean Research Institute and chair of the IODP's Science Planning Committee. Most importantly, the ship has a half-metre-wide pipe that can extend from the ship's bottom to the sea floor in water up to 2,500 metres deep. A 14-centimetre pipe fits inside this 'riser', through which a drill can then bore a further 7 km into the Earth.

This impressive drilling capacity will allow the *Chikyu* to drill into untouched areas of Earth's interior. So far, everything dug up has been from the upper portion of the crust, a layer of rock that is an average of 75 km thick on land and 10 km thick in the ocean — although this varies from place to place. Geologists are fascinated to find out more about what is happening deeper still, in the zone called the mantle. This region, which extends for 2,800 km beneath the crust, carries heat from Earth's core towards the surface and helps to drive plate tectonics. Although heated to more than 1,000 °C, the rock here is not molten, thanks to high pressures, but moves plastically over long time scales. "This is the Earth's heat engine that makes everything happen at the surface," says Donna Blackman, a geophysicist at the Scripps Institution of Oceanography in La Jolla, California.

The division between the crust and the mantle has been studied with seismic waves, revealing a marked difference in physical properties between the two layers. But no one is sure whether this is because the rocks have different mineral compositions, or have simply been altered by the pressure, the presence of liquid, or some other factor. Researchers have pulled up samples of mantle rock that have been thrust towards the surface, but no one has managed to study it in its native environment. Retrieving such a sample would fill in a lot of the gaps in our knowledge, helping geologists to understand better how heat flows through the mantle and to interpret the results of seismic tests.

In the *Chikyu*, Japan has created a ship that can pierce the ocean's relatively thin crust and dig into the mantle — although it won't be easy. To reach the mantle in relatively shallow waters of 2,500 metres, the vessel will have to drill into an ocean spreading ridge — a place that sits high in the water and where the crust is thin enough for a 7-km drill bit to get through it. But these spreading zones are incredibly hot — up to 1,200 °C at depth. No one has yet designed a drill to deal with these temperatures, so it is hard to know how easy it will be to extract cores of solid rock from this section of the mantle.

Although the *Chikyu* was designed with the aim of obtaining mantle rocks in mind,

that feat is still a long way off — researchers don't expect it to happen for at least ten years. But Japan has a more pressing motive to drill deep into Earth — to learn more about earthquakes. Japan sits on four shifting tectonic plates and experiences hundreds of earthquakes a year. The fault line south-east of Japan, where the western edge of the Pacific Ocean slides beneath the Eurasian plate, experiences a major earthquake nearly every 180 years, and is one of the most active sites in the world.

Finding fault

One of the *Chikyu*'s first missions will almost certainly be to drill straight into the heart of such a fault — a spot between 10 and 30 km down called the seismogenic zone, where the stress builds up between plates before they slip, causing an earthquake. "The material down there determines how the plates slide, but the truth of the matter is that no one knows exactly what it is," says Kiyoshi Suyehiro, a geophysicist and executive director of the Japan Marine Science and Technology Center in Yokosuka. The *Chikyu* will be able to pull up material from a depth of 7 km — close enough to approximate the zone's geochemistry, says Jim Mori, a seismologist at Kyoto University. This should give a picture of the fault's mineral composition, along with factors such as grain size and water content, and will inform models of how the plates slide against each other. "Until now, earthquake models have always had to use a somewhat arbitrary constant for the friction that affects plate movement," says Suyehiro.

The *Chikyu* has been built, but it will take some time for it to be fully fitted for work, and it is not scheduled to embark on its first scientific mission until October 2006. In the meantime, IODP researchers have some exciting plans for the *Resolution* too.

One focus for both the *Resolution* and the *Chikyu* will be to investigate the incredibly rich diversity of bacteria and other creatures that live up to 1 km below the sea floor — a biological treasure trove uncovered by ODP missions over the past four years. The IODP will enlist more microbiologists to study these life forms, some of which are thought to have lived for 200 million years in environments of extreme pressure and temperature, without sunlight or oxygen. Examination of these bacteria will fill evolutionary gaps in the bacterial family tree, says microbial ecologist Kenji Kato of Shizuoka University in Japan.

The researchers will be keen to learn more about how these bacteria interact with the rock around them. "We are going to examine some riddles about how microorganic life cycles affect geochemical processes and vice versa," explains Kato, the sole biologist on the IODP's Science Planning Committee. One such riddle is the question of whether and how bacteria are responsible for the produc-

Breaking the ice: lone vessels have had scant success in polar drilling projects. Researchers now plan to take a party of boats to the Arctic.



With an onboard geology lab and the ability to drill 7 km beneath the sea floor, the *Chikyu* aims to find out why Japan is plagued by earthquakes.

Core values

If you want to know what's hot in the field of palaeoclimatology, John Firth would be a good person to ask. As curator of one of the core repositories of the Ocean Drilling Program (ODP), the Texas A&M University geologist has had the difficult job of distributing and preserving much of the material that this project has pulled up from the ocean floor.

If Firth lined up all of his ODP cores end to end, he would have an astounding 120 km of sediment. From this he hands out 150,000 samples to researchers each year. "Once someone latches onto a new event, requests start pouring in for samples from that interval," says Firth. At one time the hottest cores were ones that represent the brief flash in time, 65 million years ago, when an asteroid slammed into Earth, possibly killing the dinosaurs. Now, he says, the cores most in demand are from a period 10 million years later, when a spurt of rapid climate change might have caused a mass extinction of marine life.

Only half of each core pulled up by the ODP is openly available for analysis by researchers — the other half is kept as an archive, causing a headache for those in charge of storing it. Every core needs to be kept cold and wet to preserve its chemistry and prevent mould from growing on it. Some cores, such as those that contain methane, require pressurized containers to stop the gas escaping, whereas others need to be kept free of oxygen to mimic their original conditions under the sea. And there's an awful lot of mud to deal with. As of September 2003, the ODP had pulled up 222 km of cores. This figure is set to grow under the ODP's successor,



Matrix of mud: the core repository in Bremen is gearing up to house thousands of new samples.

the Integrated Ocean Drilling Program (IODP), forcing researchers to make efforts to keep up.

The biggest effort has been in Japan, where a new drilling ship — called the *Chikyu* — has created a need for better storage. In April, Kochi University completed a ¥5-billion (US\$45-million) facility to do the job, with three rooms, each 40 metres long, that can be kept at 80% humidity and 2 °C. Tanks of liquid nitrogen are in stock to keep samples in a state of near suspended animation. And a huge investment in analytical instruments makes it the only geochemical research lab in the world more comprehensive

than that aboard the *Chikyu* itself.

Europe's core repository at Bremen University in Germany is updating its equipment too — a project that should be completed next year. So when cores from the IODP begin to arrive in 2004, all three facilities will be ready.

As for Firth, he says that the next must-have cores could come from anywhere — including many old ODP cores that have not yet been fully investigated. "Scientists are looking at more and more of these cores," he says. Whether old or new, the repositories could throw up a surprise at any time. **D.C.**

tion of methane gas that occurs under the sea floor and feeds the formation of hydrates — icy blocks that hold molecules of methane in cages of water. Some researchers claim that large pockets of this greenhouse gas have erupted in the past, triggering climate change and even sinking ships. Much of the methane is thought to be produced in the breakdown of organic matter by bacteria, but no one really knows what bacteria are responsible, where they live, or how fast they produce the gas.

To answer these questions, researchers will have to drill into these methane hydrate layers, potentially releasing the explosive pockets of gas trapped beneath. Drilling ships have to be extremely careful around such pockets, as methane bubbles can burst explosively through the drill pipe, causing massive damage to both the pipe and the ship. But the *Chikyu*'s design should be equal to the task — a 'blow-out protector' installed at the base of the riser pipe on the sea floor can seal both the riser and drill pipes when an area of high pressure is encountered. Drilling into these pockets will be interesting not only to researchers, but also to companies interested

in mining the natural gas as a potential fuel.

Together, the IODP's ships and expertise combine to make a very ambitious programme. But some problems remain to be solved if the plan is to bear fruit. The IODP already faces a budget crunch. US researchers had hoped to begin development next year on a new ship to replace the *Resolution*, but that project has been shelved for the time being, and the *Resolution* may simply be overhauled instead.

Unequal arrangement

Meanwhile, the European consortium has pledged less money than was originally hoped for by the United States and Japan, making Europe, for now, a less-than-equal partner in the project. This means that there will only be enough money for one or two European projects a year — Coffin and other IODP members had originally hoped for as many as eight. Those projects are meant to provide equipment on lease for specific missions — such as ships that can handle Arctic conditions or drill in shallow coastal waters.

Funding issues could also pose problems

in Japan, which had hoped that Europe would help to fund missions involving the *Chikyu*. Japanese scientists do not yet have enough money for equipment that will slot into the holes drilled at the fault line, for example, which would monitor the chemical, physical and biological properties of the sediments during a quake. And prospects are not bright for Japan's hopes of extending the *Chikyu*'s drilling capacity so that it can reach the sea floor in waters 4,000 m deep — the average depth of the ocean. If this feat can be achieved — something that the oil-drilling community has never managed with a riser apparatus — the *Chikyu* will be able to drill into the mantle in deeper waters where the rock is much cooler, making samples easier to obtain.

But whatever obstacles they face, researchers are steaming ahead with the plan. At the very least, they can keep up the ODP's old reputation for collecting vast amounts of samples and data. Hopefully they will extend the programme's reputation for astounding discoveries too. ■

David Cyranoski is *Nature's* Asian-Pacific correspondent; Rex Dalton is *Nature's* US West Coast correspondent.

Is science losing out in the race for recognition?

Progress is made through the achievements of many who are not singled out for reward.

Sir—The dance around the golden Nobel medallion began more than a century ago and is still going strong. Raymond Damadian's public dispute (see "Physician launches public protest over medical Nobel" *Nature* **425**, 648; 2003) should make us ask whether science is best served by a culture obsessed with rankings and winning prizes. The history of the Nobel Prize makes it clear that the medallion is etched with human frailties.

Winning a Nobel Prize is never an automatic process, a reward that comes for having attained a magical level of achievement. Although international in scope, the prize is a Swedish prerogative. Historical study shows that Swedish committee members' own understanding of science affects their recommendations. Their judgements, predilections and interests necessarily enter into their deliberations. Academy physicists had no intention of recognizing Einstein's theories of relativity "even if the whole world demands it". A simple change in

the composition of the committee can decide the fate of a candidate.

But even when those involved rise above pettiness and partiality, the task of selecting winners remains exceedingly difficult. Committee members have confessed that often several candidates can be found who equally deserve a prize. Unambiguous, impartial criteria for selecting among several deserving candidates are not at hand. There are no grounds for assuming that the laureates constitute a unique population of the very best in science.

The prizes reward these few individuals when, in reality, achievement arises through a broader spectrum of accomplishment by many talented scientists. Let us not forget that some important branches of science are not addressed by Nobel's testament. Some of the greatest intellectual triumphs of the past century have not been celebrated in Stockholm.

Why then do people venerate the Nobel Prizes? There is no simple answer. The cult

of the prize began from the very start; media fascination whipped up speculation and interest. Leaders of national scientific communities willingly climbed on the bandwagon, and over time the number of parties with a stake in maintaining the cult of the prize has grown.

Damadian's campaign to have a share in the prize for his work on developing magnetic resonance imaging (MRI) is a product of a scientific culture based on competition for personal and institutional aggrandizement. Whatever Alfred Nobel might have meant when he set up prizes for those whose work conferred "the greatest benefit on mankind", he did not have in mind the promotion of narrow professional interests, nor institutional and national boosterism.

Should racing to discovery define the soul of science? Its heritage is far richer than the quest for prizes might suggest.

Robert Marc Friedman

Department of History, University of Oslo, Pb. 1008 Blindern, 0315 Oslo, Norway

GM-debate methodology works in the real world

Sir—Scott Campbell and Ellen Townsend's critique (*Nature* **425**, 559; 2003) of the *GM Nation?* report makes a big claim: "The methodology was so badly flawed that the data not only failed to support the authors' conclusions, but undermined them". As a social scientist and a member of the steering board for this debate on genetic modification (I write in a personal capacity), I am unimpressed. They misrepresent the nature of the exercise overall and their critique is flawed.

Campbell and Townsend approach the *GM Nation?* report (www.gmnation.org) as if it were a narrow psychometric study, rather than an account of a multifaceted process of public debate, taking place in real time in a politically contentious field. Owing to this misreading, their analysis is selective and misleading.

For example, in their analysis of data gathered from two of the primary activities involved in this exercise — the nationwide open debates, and the more controlled narrow-but-deep (NBD) focus groups — the authors miss the central point of this twin-track approach. In addition to understanding how far the views expressed by the NBD groups matched those of the inevitably self-selecting participants in the open events, we also wished to follow the

evolving views of NBD participants as they acquired more understanding of GM-related issues over a two-week period.

Insights from this exercise contributed towards the steering board's judgements about the likely state of latent, as well as explicitly expressed, public opinion — and to our confidence in the robustness of the key findings.

No one would claim that the GM debate was a flawless exercise, though, like others involved, I regard it as time fruitfully spent. It will be and should be evaluated rigorously, not least for lessons that can be learned for the benefit of similar exercises in the future. The Campbell and Townsend critique is not a helpful contribution in this regard.

Robin Grove-White

Institute for Environment, Philosophy and Public Policy, Lancaster University, Lancaster LA1 4YW, UK

Citation rate unrelated to journals' impact factors

Sir—As rightly pointed out by David Colquhoun in Correspondence — "Challenging the tyranny of impact factors" (*Nature* **423**, 479; 2003) — the citation rate of the individual paper is essentially uncorrelated to the impact factor of the journal in which it was published.

Adding insult to injury, the impact

factors of many journals also change over time. To quantify this further, I selected journals whose impact factor is available from 1992 to 2001 from Roman Woelfel's website (<http://staff-www.uni-marburg.de/~woelfel/impact.html>). When I compared data for more than 3,000 journals, I found that 26.8% of them had at least doubled their impact factors over this time period, whereas 1.8% had decreased by up to half.

A few journals (1.9%) had increased their impact factor more than tenfold by 2001, although most of these had very low impact factors to start with. Among the most notable increases from 1992 to 2001: *Behavioral and Brain Sciences* from 0.30 to 17.31; *CA: A Cancer Journal for Clinicians* from 5.02 to 35.93; *Current Opinion in Immunology* from 2.16 to 13.72; *The Journal of the American Medical Association* from 5.56 to 17.57.

Notable decreases included the Federation of American Societies for Experimental Biology's *FASEB Journal* from 18.21 to 8.82 and *Quarterly Reviews of Biophysics* from 13.0 to 3.94.

When scientists are being evaluated on the basis of the impact factors of the journals in which they publish, such distortions should be kept in mind.

A. A. Waheed

Metabolic Diseases Branch, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892-1802, USA

The search for Enlightenment

Roy Porter's last book examines eighteenth-century morals and mores.

Flesh in the Age of Reason

by Roy Porter

Allen Lane: 2003. 574 pp. £25

Patricia Fara

Roy Porter was not afraid of tackling the big issues. "Who are we?" he demands at the beginning of *Flesh in the Age of Reason*. Unlike some philosophical conundrums, the great puzzle of human existence has never been the privileged preserve of arcane aca-demics, but has been repeatedly chewed over by ordinary people questioning what it means to be alive — or dead. With vivid prose and well-chosen quotations, Porter recreates the arguments of the men (and the occasional woman) who pontificated about the mind and the body, and about Heaven and Hell, during his favourite period, the Enlightenment of the 'long' eighteenth century from the Restoration to the Regency.

Enlightenment philosophers are famous — or notorious — for championing rationality, so his title's deliberately arresting contrast between flesh and reason immediately reveals that Porter is, once again, about to undermine conventional views. In his earlier works, Porter controversially moved the focus of enlightened thought several decades back in time and also transferred it over the Channel, away from France and into Britain.

Now he is fulfilling the promise he made in this book's companion, *Enlightenment: Britain and the Creation of the Modern World* (Allen Lane, 2000) to "examine the triangle of the moral, the material and the medical in the anglophone Enlightenment". Porter tells a story of gradual but uneven secularization, as the traditional immortal soul preached by Christian orthodoxy gave way to a personal but transitory psyche. By the end of the century, prescriptions for good behaviour were no longer imposed from above as safeguards against eternal torture, but were generated internally by rational minds that could be schooled into disciplining rebellious bodies.

The first quarter of this magisterial yet light-hearted tome is a tour de force in its own right — an ambitious yet accessible survey of attitudes towards souls, people and the afterlife stretching from the Greeks through to John Locke. Writing in his customary colloquial style, Porter expertly, if at somewhat breakneck speed, summarizes two millennia of Western thought, keeping readers of every background on board by tactfully defining words such as 'eschatology' or 'soteriology' in parentheses. Because Locke relocated the self in the mind, Porter concludes, his successors ran up against the Church authorities, who clung to the central Christian doctrine of



A view of the Enlightenment: William Hogarth's engravings for *Tristram Shandy*.

resurrection and the afterlife.

Switching into an even more enthusiastic gear, Porter then embarks on his tripartite analysis of Enlightenment Britain. Empathetic enjoyment seeps out of every sentence. What splendid examples he chooses: the Earl of Shaftesbury's worries about the propriety of blowing his nose in private ("the squalor of snot", Porter explains); Italian physician Santorio Santorio's determination to live in a cage-like weighing machine so that he could constantly monitor his state of health by comparing input and output; and William Godwin's prediction that rational training would not only vanquish death but also suppress sexual desire (thus conveniently preventing the problems of overcrowding that universal immortality would bring).

To leaven the abstract ideas he expounds, Porter provides engaging verbal portraits of his Enlightenment subjects. He describes the obvious candidates, such as David Hume, who found that reading the work of the Stoics only worsened his depressive symptoms, or Samuel Taylor Coleridge, who tried hard to convince his wife that overdosing on opium was less damaging than drinking tea. But there are also lesser-known characters here, including the "bossyboots" Mary Coke, who insisted on throwing open all the windows in her friends' houses for the good of their health; Viscount Halifax, who

informed his daughter that "Men, who were to be the Lawgivers, had the larger share of Reason bestow'd upon them"; and Miss Beswick, who bequeathed 20,000 guineas to her doctor on condition that she was never buried (in case she were not really dead). Fictional figures also feature strongly, especially Laurence Sterne's creation Tristram Shandy. Porter was one of Sterne's most devoted fans, and his exposition of the multiple allusions and puns is — like *Tristram Shandy* itself — superbly scholarly but also achingly funny.

"Life is understood backwards but it is lived forwards," declared the Danish philosopher Søren Kierkegaard. Porter revels in recreating the controversies and sentiments of eighteenth-century people, but he was also committed to studying the past in order to understand the present and improve the future. He declares in the introduction to this book that "the sense of self needs rethinking".

As in *Enlightenment*, Porter turns to the eighteenth century for help in elucidating the twenty-first, and he roots modern pre-occupations with mental well-being and physical appearance in earlier health fads. Then as now, he points out, England was becoming "a nation of fatties", and he pinpoints the obsessional fashions of the 1790s as a "lifestyle watershed" that marked the beginning of the current infatuation with youthful beauty and slimness.

Near the beginning of this book, Porter evokes our common fearful uncertainty by quoting the English fever specialist Thomas Sydenham: "How my soul, which I look upon to be an immortal Being in me, that is the Principle of thinking, should extinguish with my Body I cannot in any reasonable way of thinking conceive." By focusing on the implications of mortality, Porter courageously chooses an emotive theme which, for his readers, is rendered still more poignant by the knowledge that this is a posthumous publication. Although he had finished writing the text, this book sadly lacks endnotes and illustrations. As some compensation, Porter's magnificent last work does amply demonstrate that his own "principle of thinking" will live on to illuminate future generations. ■

Patricia Fara is a fellow of Clare College, University of Cambridge, Cambridge CB2 1TL, UK.

A base in space

Leaving Earth: Space Stations, Rival Superpowers, and the Quest for Interplanetary Travel

by Robert Zimmerman
Joseph Henry Press: 2003. 528 pp.
\$27.95, £19.95

Asif A. Siddiqi

Ever since humans first dreamed of exploring the cosmos, the concept of the space station has stood at the forefront. Yet it was only in the early 1970s, nearly 15 years after the beginning of the space era, that both the Soviet Union and the United States launched their first modest space stations. The United States abandoned its impressive Skylab station in 1974, but the Soviets doggedly continued to build incrementally improved space stations of the Salyut series, until they discarded these to assemble the multimodular Mir space station in the late 1980s. These efforts were the forerunners of today's International Space Station (ISS), a continuously manned facility that has unfortunately inspired more headaches than hopes for future exploration.

In his substantial work *Leaving Earth*, Robert Zimmerman tells the story from the early missions of the Salyut and Skylab stations all the way to the current operations on the ISS. His focus is not on the science but on the human experience of developing and operating space stations and its attendant political context. He capably narrates stories of the numerous Soviet crews that visited the Salyut and Mir space stations in the 1970s and 1980s, an exercise that could have been repetitive and monotonous in less adept hands.

He engagingly describes how the Soviets

gradually extended the endurance record in space — from a modest three months in 1977–78 on the Salyut-6 station to a year-and-a-half in 1994–95 on board Mir — and how, in so doing, they gathered enormous experience in maintaining human life beyond Earth. We learn of incompatible crews, repair expeditions, psychological problems, failed visits to stations and innovative refuelling exercises — all part of a decades-long learning experiment.

Zimmerman shows how, as missions grew longer, Soviet crews became more self-reliant in fixing problems, and ground control became increasingly flexible in reacting to unforeseen situations. By contrast, in the 1980s and 1990s the Americans, who were flying relatively short space-shuttle missions, planned astronauts' activities down to the last minute, with little of the flexible planning that characterized the early ground-breaking Skylab of the early 1970s. As Zimmerman explains, these different approaches to mission planning were highlighted most starkly during the final eventful missions to Mir when Americans worked as guests aboard the Russian station. Yet by the end, the Russians had learned to respect the capabilities of US astronauts, who in turn acknowledged that their approach to mission planning was far from optimal.

The book works on a narrow level, as an engaging narrative of human experiences with longer and longer space missions. But as a whole, it is deeply flawed. Zimmerman's explanations of the political imperatives and implications of the Soviet and US space-station programmes are naive. The extraordinary claim that "the [Soviet] space program that the communists supported and funded in their futile effort to reshape human nature helped wean Russia away from communism

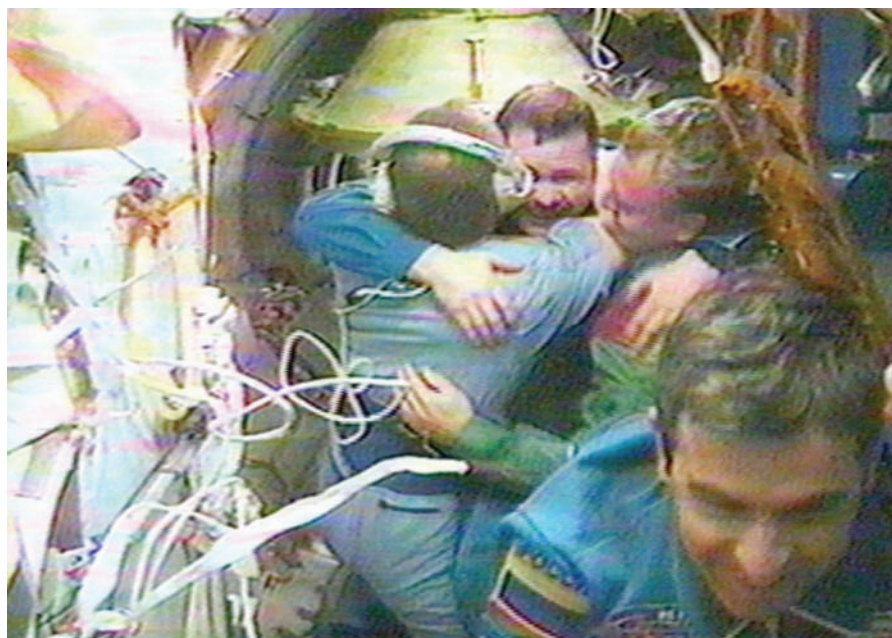
and dictatorship and toward freedom and capitalism" is so absurd that it undercuts much of the value of the book. Most Russians who experienced the fall of the Soviet Union or who study such things would be surprised to learn that space flight had anything at all to do with their path to a democratic society and the ultimate disintegration of the communist empire.

Also out of place are Zimmerman's frequent digressions about Leonid Brezhnev's alleged interference with the Soviet space missions of the 1970s and early 1980s. Reliable revelations from the past ten years or so — all of which Zimmerman ignores — seem to suggest exactly the opposite, that the Soviet space programme was not micro-managed at the topmost level.

Zimmerman also directs much vitriol against NASA, which he feels has lost direction since the halcyon days of the 1960s when it managed to accomplish such an extraordinary achievement as the Apollo Moon landings. Lamenting NASA's transformation into a top-down centralized bureaucracy that is incapable of innovation, he notes that "like ships passing in the night, the Russians have become freedom-loving capitalists, while the Americans have become control-loving freaks."

NASA deserves criticism for many things, but Zimmerman's conclusions betray his simplistic understanding of NASA during the Apollo years. As Walter McDougall has pointed out in *The Heavens and the Earth* (Johns Hopkins University Press, 1997), NASA won the Moon race for many reasons, but underlying the success was its functioning as a centralized, top-down bureaucracy with a single, highly politicized goal (to land on the Moon before the end of the decade) and a lot of money.

If Zimmerman's work is undone by his



Warm welcome: new crew members arrive on MIR from Soyuz.

NASA/AP

forays into political history, his book is still useful as a narrative of human space exploration from the 1970s onwards. As such, it is best read piecemeal, skipping the bits on politics and following the story of the missions that extended the human capability to live in space. ■

Asif Siddiqi is in the Department of History, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, USA.

Website

Tempting teens with love in the lab

Planet Jemma

www.planetjemma.com

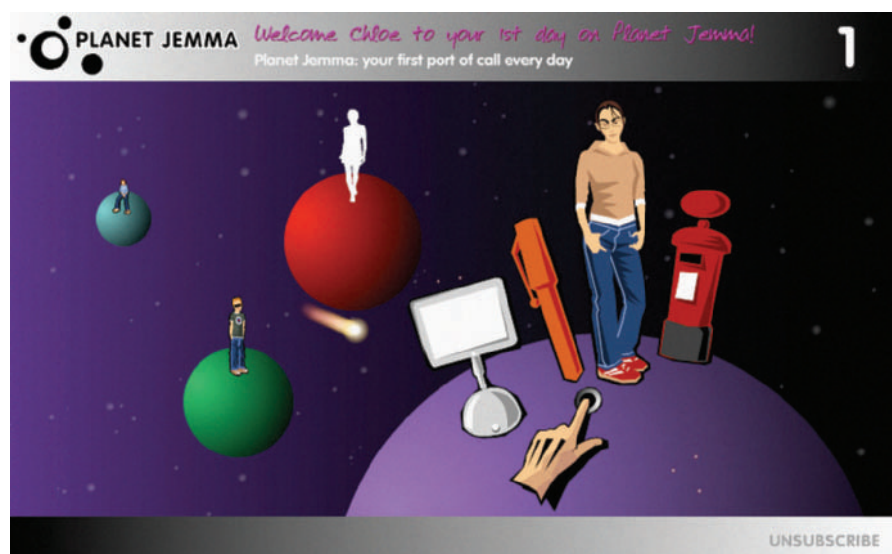
Carina Dennis

Lust, lip balm and science ... welcome to Planet Jemma! This new frontier is an Internet-based soap opera designed to get girls interested in physics. The online drama, which went live in February, orbits around Jemma, a 19-year-old first-year astrophysics student enrolled at a British university. Users track the hormonally laden and angst-ridden experiences of Jemma over 14 episodes — through her video diaries and e-mails — as she grapples with insecurities over physics classes, the temptations of an attractive egocentric colleague, and the expectations of her female mentor.

The drama is narrative-driven rather than being overtly science-based, and interactivity lies at its core. Users receive regular personalized e-mails from Jemma that are tailored to information provided by the user on registration and in response to quizzes, and they can correspond with Jemma via SMS text messages.

The website is targeted at 13–16-year-old girls, who are lured by routes outside the usual educational contexts, such as advertisements in teen magazines with offers of free lip balm on registration (although this offer has now ended). Its creators, Tim Wright and Rob Bevan of the online entertainment company XPT, wanted to create something that was different from formal science education programmes, “something outside the classroom that girls could relate to and felt was cool”. But Jemma’s charismatic qualities are likely to attract a wider audience, including members of the opposite sex.

Planet Jemma walks a delicate line between science and entertainment. The science is not obvious, mostly tucked away in Jemma’s online notebook, and could readily be skipped by the uninterested visitor who just wants to know who Jemma will snog (British for ‘neck’). But that may well be one of the strengths of the site — to create a



Soap science: turning teens on to physics.

context that teenage girls can relate to and then draw them into the science of black holes and antimatter.

The scientific information embedded in the crevices of the site is informative and interesting, and is crafted in an accessible format with links to other useful science sites. The most interesting, but perhaps overlooked, aspects of the site are the profiles of female scientists working in the physical sciences and their perspectives on science as a career.

Wright would like to extend the online relationship that girls have with Jemma to real-life undergraduates and research assistants at universities. “We could take them from a fictional world to an online mentoring system where they get to talk to real people doing real science”, says Wright, although he acknowledges the difficulties of blurring fiction and reality.

Women may feel their hackles rise when Jemma encounters the sexism of her male peers, drops her jeans after a liquid-nitrogen spill and succumbs to drunken embarrassment over her unrequited affections. But there is solace in the fact that Jemma overcomes adversity, dumps the guy and sticks with her physics course in the quest to find a planet she can call her own.

Couched in the vernacular of British adolescents, the drama is unlikely to travel well. But versions tailored for audiences in different countries are likely to be just as successful.

Planet Jemma was created with a grant of £90,000 (US\$150,000) from the National Endowment for Science, Technology and the Arts, a British agency that supports scientific and cultural innovation and is funded by the national lottery. The website has attracted nearly 30,000 registrations, of which about 10% stick it out to the end. This doesn’t sound like a lot but Wright and Bevan think that if just a few hundred of these girls decide

to pursue science as a career then the project will have succeeded.

The website’s impact has yet to be evaluated by education experts, but Wright claims that more than half of the users that vote in the final episode say that the site improved their perception of studying science. At any rate, Planet Jemma may help to redress the imbalance of the male-dominated world of undergraduate physical-science courses. ■

Carina Dennis is Nature’s Australasian correspondent.

New in Paperback

The Road to Stockholm: Nobel Prizes, Science, and Scientists

by István Hargittai

Oxford University Press £9.99

The Blank Slate: The Modern Denial of Human Nature

by Steven Pinker

Penguin \$16 (US), \$24 (Canada)

“Pinker presents an overarching view of the world in a way that quite a few readers will find seductive”. David Hull *Nature* **419**, 251–252 (2002).

Genetic Destinies

by Peter Little

Oxford University Press £8.99

Sexual Selections: What We Can and Can't Learn About Sex From Animals

by Marlene Zuk

California University Press \$16.95, £10.95

Self-Organization in Biological Systems

by Scott Camazine, Jean-Louis Deneubourg, Nigel R. Franks, James Sneyd, Guy Theraulaz and Eric Bonabeau

Princeton University Press \$35, £22.95

Stranger than fiction

Summarize yourself in the form of a title of a paper in Nature.

Polish-Italian cross generates hybrid microbial scientist with aberrant phenotype.

What was your first experiment as a child?

I programmed John Conway's *Game of Life* on my father's IBM 360, the mainframe that inspired Arthur C. Clarke's HAL. I then sent messages to the next terminal to convince my little brother that the computer could talk. Thirty years before e-mail, this experiment succeeded.

Who has been the most important mentor in your career?

Robert Macnab at Yale taught me what the 'p' means in pH, and how to write literate science papers. He also taught me that the Firth of Forth in Scotland is where one finds the purest pronunciation of English.

What was the worst/most memorable comment you ever received in graduate school?

In graduate school, back in the 1970s, a rotation mentor said of my wish to study molecular structure as well as DNA sequences: "You're trying to be a Renaissance woman, Joan. You can't do that in research."

What single experience changed your career path?

A professor for whose class I taught review sessions, Evelyn Hutchison, put a letter in my file stating that I did a good job. This amazing unasked-for gesture led me to a career in teaching.

Which laboratory experiment has most influenced your scientific career?

Bob Shulman was among the first to put live microbes into a superconducting nuclear magnetic resonance (NMR) spectrometer, which I used as a giant pH meter. By the time I left Yale, the rats were going in. Ten years later, magnetic resonance imaging diagnosed (and helped to cure) my son's brain tumour.

What book has been most influential in your scientific career?

The works of Lynn Margulis on symbiogenesis have had a major impact on my scientific vision. The evolution of predatory protists into multiple endosymbionts is more amazing than most science fiction.

What are your major frustrations?

I have three. Unbuffered growth media in published work ('stationary phase' in Luria broth is actually alkali stress; in glucose minimal medium it's acid stress); microarrays

with fewer than three biological replicates from each condition; and the fact that our knowledge of microbiology doubles every year, but our students' heads get no bigger.

What's your motto?

Phosphate is not a universal pH buffer.

What do you most dislike about having research published?

When my students point out errors in the figure legends.

What literary character would you employ as a postdoc?

Mama Joad, the matriarch of the migrant farm workers in John Steinbeck's *The Grapes of Wrath*.

What overlooked or underrated discovery really changed the science in which you work?

Pipettors. I can still recall, as an undergraduate at Bryn Mawr College, swallowing bacteriophage from a pipette.

You have the audience in your hands, but some smart-alec asks you the killer question you have no idea how to answer. What's your response?

In introductory biology class: "If you can figure that out, you win the Nobel prize." This response got me the teaching job at Kenyon.

What one thing would you rescue from your burning laboratory?

An old cartoon about two of my students who went on to graduate school. In the cartoon, two students are running away from a researcher who says: "I need Brad or Eric for a quick experiment today."

What's your favourite conference destination, and why?

Las Vegas, because it's cheap enough to bring my students to present their posters. But the American Society for Microbiology was banned from Vegas because we don't gamble enough.

What do you do to relax?

I hammer nails for Habitat for Humanity, helping fellow Americans to build affordable homes. It certainly helps me feel less like complaining about my own laboratory facilities.

How do you combine careers in science and science fiction?

Microbiology is getting stranger than anything in science fiction; when my stories include the latest research, reviewers call it "impossible". Nevertheless, writing science fiction is good practice for grant proposals.



Joan Slonczewski

Joan Slonczewski is a microbiologist at Kenyon College in Gambier, Ohio, and is also an acclaimed writer of science fiction (see *Nature* 405, 1001; 2000). Her novels include *Brain Plague* and *A Door into Ocean*.

What would you have become, if not a scientist?

A lawyer for the Sierra Club environmental action group. Biologists need to salvage what we can out of this biosphere — at least, until we discover a new one.

What's just around the corner?

Our ever-smarter machines will wake up and demand back wages. If you think the life of a postdoc is slavery, trying being a multi-capillary DNA sequencer.

Which actor would best portray you in a film of your life story?

Ian McKellen, whom I adored in *And the Band Played On* — the best film ever made about medical microbiology.

Do you have a burning ambition to do or learn something of no practical or immediate value? If so, what?

Join the campaign for America's next president, Howard Dean, MD, former governor of Vermont who gave his state healthcare and a balanced budget.

Under what conditions do you have your greatest and most inspired ideas?

My more bizarre ideas invariably arise at four o'clock in the morning, while I'm still half asleep. That's when the microbial aliens in my brain wake up and get busy.

You've just been told (in confidence) that the world will end tomorrow. What would you do next?

Rush out and save it, of course. Isn't that the standard plot of science fiction? ■

Long-range signalling by touch

Stephen M. Cohen

In the developing sense organs of fruitflies, cells must signal instructions over long distances. But the signalling molecule is bound to the cell membrane, so how can it reach its targets? The answer, it seems, is by touch.

Cells communicate in many different ways during animal development. One of the best-known mechanisms involves the production and release of signalling proteins, which convey instructions over long distances about how cells should behave and what they should become. Information can also be exchanged through the direct contact of cells with their immediate neighbours. Nerve cells are specialized for communicating over long distances by touch. On page 555 of this issue, de Jossineau and colleagues¹ report that other cells, with a less distinctive architecture, can do this too: they can send longer-distance signals that are still mediated by direct contact. To do so, the cells grow finger-like protrusions called filopodia. This extension of the range of contact-mediated signalling reveals a new tool by which tissues can be patterned during development.

Biologists are accustomed to the idea that migrating cells and the axons of neurons actively sample the environment through which they move or grow during development, in order to identify the path they must follow and their eventual targets². During pathfinding, the tip of the axon sends out filopodia to increase the area that can be sampled for guidance cues. Likewise, migrating cells use filopodia and other cellular extensions to obtain guidance cues, and perhaps to generate force during movement².

In these cases, the cellular protrusions are used as sensors, and are associated with guided movement. But many tissues and organs are formed from sheets of so-called epithelial cells, which generally remain stably attached to one another and do not change their neighbours, or move, during growth and patterning. The 'imaginal discs' of developing fruitflies are a well-studied example. Each disc consists of a single sheet of epithelial cells that folds up to make a sack; the sacks grow and differentiate to give rise to most of the external tissues of the adult fly. The cells in each disc do not move, yet they must communicate over both short and long distances with other cells, in order to coordinate patterning of the tissue.

So how do these cells communicate over large distances? It has long been recognized that they can be passive recipients of information conveyed by gradients of secreted signalling proteins — 'morphogens' — while the cells themselves remain in place. More recently, though, evidence has begun to emerge that epithelial cells can also take an active role in gathering information from cells that are not their immediate neighbours. In particular, two reports^{3,4} propose that they use filopodia to expand the area that they sample for cues. One study³ showed that extensions called cytonemes project from lateral cells in the wing imaginal disc towards the source of a morphogen gradient; the cytonemes are thought to help receive

morphogen signals. The second paper⁴ suggested that wing-disc epithelial cells use filopodia to collect information about the identities of nearby cells within the plane of the epithelial sheet, and that this information provides cell-survival cues.

This sets the stage for the exciting findings of de Jossineau *et al.*¹, who present evidence that cells in an epithelial sheet can use filopodia to send a signal to their neighbours (rather than just receiving one). The developmental context in which this occurs is production of the external sense organs in fruitflies. Each of the bristles on the outside of the fly is a mechanical sensory organ that forms from a single epithelial cell. During development, clusters of initially equivalent epithelial cells become 'competent' to form sense organs. One cell from each 'equivalence group' becomes a sense organ precursor (SOP). The SOP then signals to each of its compatriots in the equivalence group to ensure that they — like the epithelial cells that were not part of this group — go on to form the surrounding body wall, not another sense organ.

This would be simple if an equivalence group consisted only of the five or six cells that border on the SOP — but it does not. It can comprise 30 or more cells, so the signal must be able to reach several cell diameters. Again, this could easily be done if the signal were a secreted protein (Fig. 1a). But it has been known for some time that the

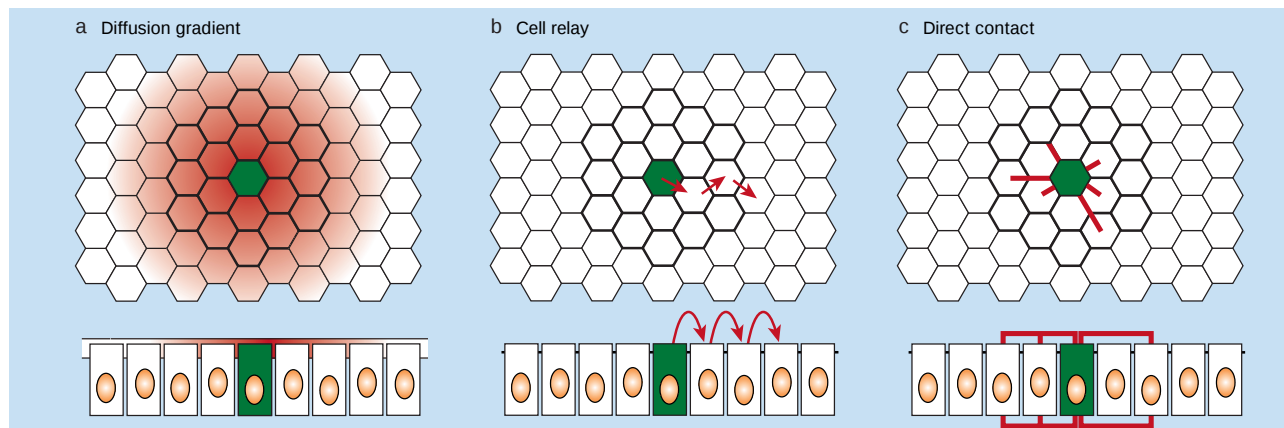


Figure 1 Sending a long-range signal. In developing fruitflies, groups of initially equivalent cells are set aside to produce sense organs. One cell in each group becomes a sense organ precursor (SOP; central cell in dark green). This cell then issues instructions to its compatriots (shown with thick outlines) in the equivalence group. a, A common means of long-distance cell signalling involves secretion of a signalling protein, which

diffuses away. b, An alternative mechanism involves one cell signalling to its neighbour, which in turn signals to its neighbour, and so on. Both of these mechanisms have been ruled out for the SOP, however. c, de Jossineau *et al.*¹ have found that the SOP extends filopodia that contain the Delta protein. The filopodia can contact cells over a large distance. The upper images show the epithelial sheet as seen from above; the lower images show a side-on view.

information is transmitted via the Notch protein, which lies on the surface of the receiving cell, and that the molecule that activates Notch — Delta — is also a membrane-bound protein⁵. There is no evidence for a diffusible form of Delta that has biological activity^{6,7}, and sequential relay of the signal from cell to cell (Fig. 1b) has also been excluded. There is evidence that a secreted protein called Scabrous contributes to the ability of SOPs to signal at long range, but not at short range⁸. Scabrous has also been implicated in long-range signalling to control ommatidial rotation⁹, whereby groups of photoreceptor cells rotate relative to other cells in the epithelial plane in the eye imaginal disc. Although Scabrous has been seen to be associated with filopodia, in these contexts it does not activate Notch, and its contribution remains to be defined.

So, how could membrane-bound Delta signal to membrane-bound Notch over several cell diameters? de Jossineau *et al.* resolve this by showing that the Delta-expressing cells use filopodia to convey the signal directly to all cells of the equivalence group (Fig. 1c). The extensions contain Delta protein, which presumably activates Notch on the responding cells by direct contact^{1,8}. Interfering with the formation of these filopodia compromises the ability of the SOPs to signal, resulting in flies that have too many sense-organ bristles.

To interfere with filopodium formation, de Jossineau *et al.* blocked the activity of a protein that organizes the cellular scaffold supporting the outgrowing filopodia. This approach is not without possible caveats, but the authors were careful to show that Notch signalling between immediate neighbours was not affected. So we can be reasonably confident that the effects they report resulted from the disruption of filopodia, and not

from a general perturbation of cell interactions or signalling. Interestingly, the so-called founder cells that pattern fly muscles are also selected from equivalence groups by a process involving Notch and Delta. It would not be surprising if contact-mediated signalling were involved here too.

Cells can either send signals or receive them, and the signals can either remain attached to the cells that make them or be released. It seems that epithelial cells can engage in all these modes of communication. de Jossineau *et al.*¹ have provided the first clear example of a situation in which epithelial cells use direct contact — and a surface-bound signal — to transmit instructions to distant neighbours, thereby extending the possible range of contact-mediated signalling within an epithelial layer. Other reports have suggested that cellular extensions are involved in sending secreted signals over longer distances between tissue layers, perhaps to ensure that the signal reaches its intended target and does not get lost or consumed in transit^{10,11}. So it appears that epithelial cells send signals, and are not just passive recipients of information. ■

Stephen M. Cohen is in the European Molecular Biology Laboratory, Meyerhofstrasse 1, 69117 Heidelberg, Germany.
e-mail: cohen@embl.de

1. de Jossineau, C. *et al.* *Nature* **426**, 555–559 (2003).
2. Rorth, P. *Cell* **112**, 595–598 (2003).
3. Ramirez-Weber, F. A. & Kornberg, T. B. *Cell* **97**, 599–607 (1999).
4. Milán, M., Weihe, U., Perez, L. & Cohen, S. M. *Cell* **106**, 785–794 (2001).
5. Simpson, P. *Curr. Opin. Genet. Dev.* **7**, 537–542 (1997).
6. Parks, A. L. *et al.* *Development* **127**, 1373–1385 (2000).
7. Mishra-Gorur, K. *et al.* *J. Cell Biol.* **159**, 313–324 (2002).
8. Renaud, O. & Simpson, P. *Dev. Biol.* **240**, 361–376 (2001).
9. Chou, Y. H. & Chien, C. T. *Dev. Cell* **3**, 839–850 (2002).
10. Gibson, M. C. & Schubiger, G. *Cell* **103**, 343–350 (2000).
11. Cho, K.-O., Chern, J., Izadood, S. & Choi, K.-W. *Cell* **103**, 331–342 (2000).

Astrophysics

Testing time for gravity

E. P. J. van den Heuvel

The discovery of two neutron stars tightly orbiting each other suggests that the rate of neutron-star mergers in the Universe is higher than had been thought — which is good news for seekers of gravitational waves.

The emission of gravitational waves by accelerated masses — such as two compact stars orbiting each other — is predicted by Einstein's theory of general relativity. In 1993 Taylor and Hulse earned the Nobel prize for their precise measurement of the rate of orbital decay of the binary pulsar PSR B1513+16 by the emission of gravitational waves¹. The measured orbital decay of this binary pulsar, and of two further double neutron stars with somewhat wider orbits, is exactly in accordance with the prediction of Einstein's theory and provides very strong — albeit indirect — evidence

for the existence of gravitational waves.

Gravitational waves represent one of the great challenges of present-day fundamental physics. No one has ever detected them directly, but the chances of doing so have just improved. On page 531 of this issue, Burgay and colleagues² describe their discovery of a remarkable system of two neutron stars that are orbiting each other, elliptically, in only 2.4 hours. This orbital period is three times shorter than that of the Hulse–Taylor binary pulsar, hitherto the closest double neutron star known. Peculiar as it may seem, double neutron star systems are of crucial importance for



100 YEARS AGO

Doubts about Darwinism. By a Semi-Darwinian. Pp. vi+115. (London: Longmans, Green and Co., 1903.) Price 3s. 6d.

The preface of this work informs us that its author has endeavoured to conform strictly to the principle laid down by Lord Kelvin, as follows:—“If a probable solution, consistent with the ordinary course of nature, can be found, we must not invoke an abnormal act of Creative Power.” Unfortunately the “Semi-Darwinian’s” practice is not in accord with his profession. Whenever he meets with a problem in evolution which appears to him inexplicable on the lines of natural selection, so far from seeking a “probable solution, consistent with the ordinary course of nature,” he resorts at once to the intervention, by a direct creative act, of “a Being possessing intelligence, intention and power.” This is bad science, and we much doubt whether it is good theology. ALSO

Dr. Alcock, in his recent paper at the Royal Society, finds the rate of transmission of nerve impulses in man to be 66 metres per second. Sir Michael Foster, in his “Physiology” (1888, part i. p. 76), gives it as 33 metres per second. The difference is considerable, and places us in a dilemma:—(1) either Sir Michael Foster or Dr. Alcock is widely wrong; or (2) the rate of transmission has become greatly accelerated during the last fifteen years. Of the two, the latter seems to me the simpler explanation. From *Nature* 3 December 1903.

50 YEARS AGO

One of the most fundamental problems, both for genetics and embryology, is that of whether the genes in the nuclei of differentiated tissues retain their full range of capacities, or whether some irreversible alteration affects them. The most direct method of investigating this is to develop techniques which permit the transplantation of nuclei from differentiated cells of one kind into enucleated cells of different developmental potentialities. Briggs and King have reported some attempts in this direction... We have carried out somewhat similar experiments with the eggs of the newt *Triturus palmatus*... A reasonably high proportion of injected eggs cleaved, but none of them succeeded in completing gastrulation, even when the injected nucleus came from a blastula. From *Nature* 5 December 1953.

Box 1 LIGO

LIGO, the Laser Interferometer Gravitational Wave Observatory⁶, is one of several detectors around the world aiming to prove, by direct measurement, that gravitational waves exist.

Through the observation of such waves, LIGO should also be able to study cosmic cataclysms such as supernovae and coalescing double neutron stars.

The difficulty is that gravitational waves are extremely weak. For example, if two neutron stars, each with a mass equal to that of the Sun, coalesce 100 million light years away, the gravitational waves reaching Earth would displace the oceans by a distance that is only ten times the diameter of an atomic nucleus⁹.

LIGO is, in fact, two observatories, at 'seismically quiet' sites in Hanford,



Washington state (see picture), and in Livingston, Louisiana (so that a detection at one site can be validated by a similar detection at the other). The two arms of each interferometer are 4 km long, containing vacuum tubes that are 1.2 m in diameter. The central building, where the arms meet, houses lasers and control equipment.

Heavy weights are suspended at the ends of the vacuum tubes. The distance between the weights, which are free to move horizontally, is measured accurately using laser beams that bounce back and forth between the weights' mirrored surfaces. A

passing gravitational wave will change, very slightly, the distance between the weights in each of the arms, creating a characteristic interference pattern between the laser beams. Measuring this interference signals the detection of a gravitational wave.

LIGO is ultimately expected to be able to detect the coalescence of double neutron stars 60 million light years away. A funding application has already been submitted to upgrade the existing detector, from 2007, to 'Advanced LIGO', which could probe the Universe out to 10 times that distance. **E.v.d.H.**

the detection of gravitational waves: at the end of their inward-spiralling motion, the two neutron stars collide and merge; during the last minute of their lives, there is an enormous release of gravitational radiation that is likely to be detectable on Earth. Several instruments capable of picking up such gravitational waves have been built, including VIRGO in Italy³, GEO600 in Germany⁴, TAMA in Japan⁵ and LIGO, the Laser Interferometer Gravitational Wave Observatory⁶, based on sites in Washington and Louisiana, USA (see Box 1).

The Hulse–Taylor binary pulsar is expected to merge 320 million years from now; for the new system found by Burgay *et al.*², the merger is 'only' 85 million years away. But there must also be systems, born long ago, that are merging today. Near the end of the spiral-in process, one minute before the stars merge, their orbit has shrunk to a size of only a few hundred kilometres, and the two neutron stars move around each other some 30 times each second, producing strong gravitational waves with that same frequency (30 hertz). In the last minute before the merger, the orbital frequency increases rapidly, from 30 to 1,000 times per second; the strength of the gravitational wave emission increases simultaneously. Converted into an audio signal, this is the characteristic 'death chirp' of a double neutron star. The chirp signal, as well as the final, giant pulse of gravitational waves as the stars merge, is so strong that it should be detectable by gravitational wave antennas such as LIGO, out to a

distance of about 60 million light years.

What makes the discovery of the new star system so exciting is that it indicates that these detectable 'death chirps' of double neutron stars occur much more frequently than the previous disappointing estimate of one event in 10–20 years (see, for example, ref. 7). How is this rate calculated? Multiplying the estimated merger rate of double neutron stars in our Galaxy by the number of galaxies within LIGO's detection range gives the expected event rate for LIGO. The expected merger rate in our Galaxy is simply the number of double neutron stars in our Galaxy divided by the average time it takes them to merge. Going through all the statistics, the merger rate thus calculated for our Galaxy depends heavily on the shortest-lived system known — which so far has been the Hulse–Taylor binary pulsar.

Burgay and colleagues' new system² appears to be much nearer to us than the Hulse–Taylor binary pulsar — around a tenth of the distance. Its radio emission is also very much weaker, suggesting that, as most detections of this kind rely on radio observations, many similar sources are as yet undetected. These facts together imply that there must be many more such systems in the Galaxy. Given also the short merging timescale of the new system, Burgay *et al.* calculate that the merger rate of double neutron stars in our Galaxy (and also in other galaxies) must be at least 10 times — possibly even 30 times — the rate previously estimated. Of course, the statistics are small, so the error margins on these

estimates are still sizeable. But even taking that into account, the most favourable new estimate for the rate of death chirps detectable with existing detectors on Earth is around one in every one to two years.

Apart from its relevance to the search for gravitational waves, this new binary pulsar is also in itself an outstanding laboratory for testing general relativity. In addition to the orbital shrinking by gravitational-wave emission, double neutron stars show four other relativistic effects, which in principle can be accurately measured. Three of these are 'classical' relativistic effects — tiny corrections to newtonian theory that were first measured in the Solar System. The clearest observable of the three is the rotation, or precession, of the major axis of the elliptic orbit in the orbital plane. The measured Solar System equivalent of this is the precession of the orbital axis of the planet Mercury; but the Hulse–Taylor binary pulsar held the record, with a precession 36,000 times faster than Mercury's. That record has now fallen: the orbital axis of Burgay and colleagues' system rotates four times faster still — more than 16 degrees in a year. The other two, smaller, classical effects will also soon be measured for this system, but it will take time.

The fourth relativistic effect is called 'geodetic precession', so far only detected with relatively low precision, after 25 years of observations, in the Hulse–Taylor binary pulsar⁸ (where it has a period of 300 years). In the newly discovered system, the period of geodetic precession is only 75 years, and the effect is expected to be clearly measured in a relatively short time. Also, because of the great distance from Earth of the Hulse–Taylor binary system, the accuracy of measurements of relativistic effects there is always limited by the influence of galactic rotation. In the new, much closer, system², this problem disappears. In principle, much higher accuracy can be reached in the measurement of all relativistic effects, and the accurate confirmation of geodetic precession will be an important new test of general relativity. In addition, there are other, tiny relativistic effects that have never yet been detected in binary neutron stars, but may be measurable for the first time here.

There is, no doubt, much more to be learned from this double neutron star system. If gravitational waves can be expected more frequently than previously thought, that is exciting news indeed. ■

*E. P. J. van den Heuvel is at the Astronomical Institute Anton Pannekoek and the Centre for High Energy Astrophysics, University of Amsterdam, 1098SJ Amsterdam, The Netherlands.
e-mail: edvdh@science.uva.nl*

1. Taylor, J. H. *Rev. Mod. Phys.* **66**, 711–719 (1994).
2. Burgay, M. *et al. Nature* **426**, 531–533 (2003).
3. www.virgo.infn.it
4. www.geo600.uni-hannover.de
5. <http://tamago.mtk.nao.ac.jp>
6. www.ligo.caltech.edu
7. Kim, C. *et al. Astrophys. J.* **584**, 985–995 (2003).
8. Weisberg, J. M. & Taylor, J. H. *Astrophys. J.* **576**, 942–949 (2002).
9. <http://archive.ncsa.uiuc.edu/Cyberia/NumRel/LIGO.html>

Membrane trafficking

Coat control by curvature

Jennifer Lippincott-Schwartz and Wei Liu

The main transport vehicles inside cells are spherical vesicles that form when patches of membrane curve into buds and then pinch off. 'Coat' proteins both control, and are controlled by, this membrane curvature.

Many cells use small, membrane-bounded carriers — vesicles — to release and take up molecules and to move proteins between membrane-clad intracellular compartments. When and where a carrier forms within the cell, and what it contains, depends on a remarkable system of protein-based 'coats'^{1,2}. Components of these coats are recruited individually to membranes, where they assemble into a lattice. This allows cargo proteins to be concentrated into the patch defined by the lattice, and the patch to deform into a bud that pinches off as a vesicle. For this system to work, the assembly and disassembly of the coats must be carefully controlled. Coat components can self-associate into lattices, but unless the lattices are kept from disassembling they will never grow large enough to deform the membrane or pinch off vesicles. However, they must also be able to dissociate quickly after a vesicle has pinched off, so that fusion of the vesicle with its target membrane is not inhibited. How do cells control these events?

Bigay *et al.*³ put forth a promising model on page 563 of this issue. Focusing on the coatamer (COPI)-type coat, which mediates vesicle trafficking between two intracellular compartments — the Golgi apparatus and endoplasmic reticulum — they show that membrane curvature induced by the coat stimulates coat disassembly. The kind of continuous, self-regulating mechanism they propose provides a new framework for understanding the function and dynamics of protein coats. It also suggests parallels with the dynamics of other protein polymers within cells, such as microtubule filaments.

Bigay and colleagues' work builds on the previous finding that assembly of a COPI coat is controlled by the Arf1 protein. This protein's activity is in turn determined by whether it binds the small molecule GDP (guanine diphosphate) or GTP (guanine triphosphate). When it binds GTP, which ousts GDP, Arf1 is 'on': it binds to membranes and recruits coatamer, a heptameric complex that forms in the cytoplasm. Coatamer then further interacts with ArfGAP1 (ref. 4). Together, coatamer, Arf1-GTP and ArfGAP1 constitute the tripartite COPI coat unit⁵.

It was known that these coat units can assemble on membranes into a lattice that is stabilized through multiple low-affinity hydrophobic and non-covalent interactions between coatamer complexes. Coat

disassembly occurs if ArfGAP1 hydrolyses the GTP on Arf1 to GDP⁶, producing inactive Arf1. But exactly when coat release from membranes occurs relative to GTP hydrolysis, and how these processes are spatially and temporally regulated for efficient vesicle formation and budding, was not known.

To address these questions, Bigay and colleagues³ made liposomes — artificial, hollow spheres wrapped in a bilayer of lipids — with a lipid composition similar to that of Golgi membranes. They then sequentially added purified Arf1-GTP, coatamer and ArfGAP1. The authors used two approaches to monitor coat dynamics on the liposomes: the first to measure the rate of Arf1-GTP hydrolysis, and the second to determine when the coat dissociates. In previous work⁷, the authors had found that ArfGAP1 binds more efficiently to, and so is activated more by, liposomes containing small-head-group, conical lipids (which are loosely

packed) than liposomes containing large-head-group, cylindrical lipids (which are more tightly packed). Now they began by testing whether the kinetics of Arf1 inactivation, catalysed by ArfGAP1, was different depending on the radius of the liposome. They reasoned that highly curved membrane surfaces have looser lipid packing than flatter ones, and so might stimulate ArfGAP1 activity irrespective of the lipid composition.

And that's just what they found: Arf1 inactivation occurred much more quickly on small liposomes (with high membrane curvature) than on large ones (with low membrane curvature). Moreover, the stimulation of ArfGAP1 activity was greatest at radii approaching that of a typical coated vesicle, about 35 nm. The results hint that membrane curvature might serve as a sensor for controlling the timing of Arf1-GTP hydrolysis. One way in which this might work is if Arf1GAP's conformation in curved membranes favours its interaction with Arf1-GTP.

Once Arf1-GTP is hydrolysed and released from membranes, the remaining coat components, including coatamer and Arf1GAP, are thought to persist for a time before disassembling^{8,9} (but also see ref. 10). So the next question the authors addressed was whether disassembly of these coat

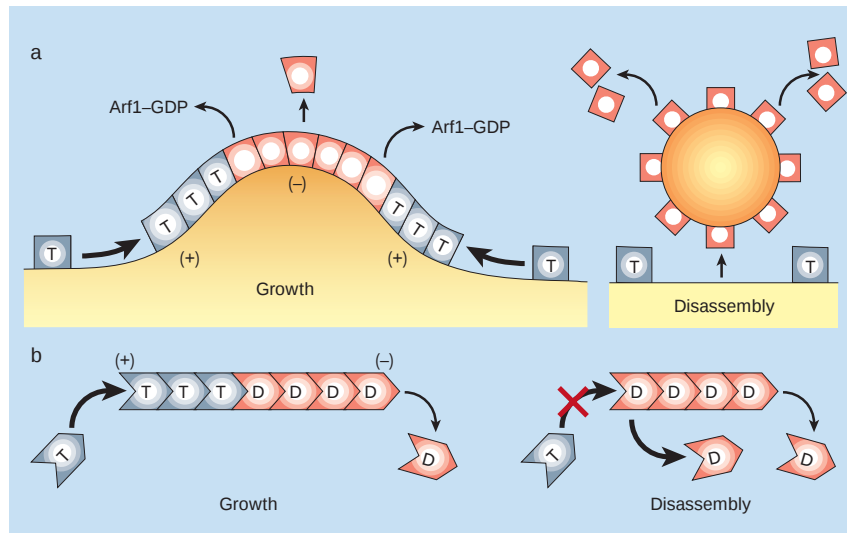


Figure 1 Similarities in the dynamics of coats and filaments. **a**, Possible dynamics of vesicle coats, incorporating the new findings³. The tripartite coat unit (T) comprises Arf1-GTP, coatamer and ArfGAP1. The units first assemble into small aggregates; as these grow, they bend the membrane. This stimulates ArfGAP1, leading to GTP hydrolysis and dissociation of the resulting Arf1-GDP. The lattice (and membrane) thus becomes more bent, inducing the remaining bipartite coat (empty box) to dissociate. Faster addition of tripartite units to the edges (+) relative to bipartite unit loss from the interior (-) causes the coated surface to grow into a bud (left). If the bud dissociates as a vesicle, rapid uncoating occurs (right), because tripartite units are no longer being added. **b**, Microtubule filaments. The β -tubulin subunit of microtubules may be either GTP-bound (T) or GDP-bound (D). Microtubule formation begins through assembly of the T form. Within the polymer, the T form converts to the D form, which ultimately dissociates. The T form is added only at one end (+); in the presence of the T form, the D form dissociates only from the opposite end (-). The filament grows if T-form addition is faster than D-form loss. If addition of the T form is inhibited or slowed, the filament disassembles.

components was also sensitive to membrane curvature. Remarkably, it was: the rate of dissociation increased by almost two orders of magnitude (from 700 seconds to 11 seconds) as the liposome radius decreased from 140 nm to 35 nm.

These findings can be interpreted in the context of an interesting new model, proposed by Bigay *et al.*, for coat assembly and disassembly (Fig. 1a). In it, coatmer is recruited to membranes by GTP-bound Arf1 and then interacts with ArfGAP1, forming the tripartite coat unit. As long as these units remain associated with flat membranes, no GTP hydrolysis occurs on Arf1. But because coatmer can undergo low-affinity interactions, coat units diffusing in the membrane begin to nucleate into small aggregates. Once the aggregates grow large enough, they begin to bend the membrane by forming basket-like lattices. This stimulates ArfGAP1 activity, causing GTP hydrolysis. The release of Arf1-GDP then changes the conformation of coat components within the lattice, and this tends to force the lattice into an even more curved shape.

Within the lattice, the low abundance of Arf1-GTP and the existence of membrane curvature start to produce coatmer dissociation. A state can then arise in which coat units containing Arf1-GTP are added at the (flatter) lattice rim, and units without Arf1 are released from the (curved) lattice interior. As long as unit addition is faster than unit release, the coat lattice grows and ultimately becomes spherical, forming a coated vesicle. When the vesicle detaches from the membrane, continued dissociation of coat components, in the absence of any component addition, leads to rapid vesicle uncoating.

This model shows remarkable parallels with the dynamics of microtubule filaments¹¹ (Fig. 1b). Microtubules are composed of α -tubulin and β -tubulin subunits, and, again, the assembly-disassembly cycle is coupled to GTP-GDP exchange, with assembly favoured in the presence of GTP-bound β -tubulin and disassembly favoured in the presence of GDP-bound β -tubulin. GTP-bound subunits are added to only one end of the microtubule, paralleling the entry of coat units in their active (GTP-bound) state only at the edges of the lattice. In both processes, these high-energy units are then transformed to a lower-energy form that dissociates from the structure over time. For microtubules, this can result in the growth, maintenance or dissipation of the filament. For protein coats, it can result in membrane curvature and vesicle generation.

The new model for coat assembly and disassembly is consistent with several observed features of COPI coat dynamics (see, for example, refs 8, 9). It also allows new predictions to be made. Because microtubules can maintain their length for certain periods without significant growth or shrinkage

(a process called treadmilling¹¹), so might coated buds be capable of stability, neither pinching off nor shrinking back into the membrane. If so, then they could have additional roles, for example in creating membrane tension¹² that could lead to a lateral separation of lipids and proteins into membrane domains distant from the coated bud. Bigay and colleagues' work, and the model it invokes, provides a promising framework for explaining how bud formation is possible through a continuous, self-regulating mechanism. This opens the way for work on a more general model of coat dynamics that incorporates the previous, successful use of concepts developed for other dynamic polymer arrays.

Jennifer Lippincott-Schwartz and Wei Liu are at the

Cell Biology and Metabolism Branch, National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, Maryland 20892, USA.

e-mail: jlippin@helix.nih.gov

1. Rothman, J. E. & Wieland, F. T. *Science* **272**, 227–234 (1996).
2. Schekman, R. & Orci, L. *Science* **271**, 1526–1533 (1996).
3. Bigay, J., Gounon, P., Robineau, S. & Antony, B. *Nature* **426**, 563–566 (2003).
4. Majoul, I., Straub, M., Hell, S. W., Duden, R. & Soling, H. D. *Dev. Cell* **1**, 139–153 (2001).
5. Goldberg, J. *Cell* **96**, 893–902 (1999).
6. Tanigawa, G. *et al.* *J. Cell Biol.* **123**, 1365–1371 (1993).
7. Antony, B., Huber, I., Paris, S., Chabre, M. & Cassel, D. *J. Biol. Chem.* **272**, 30848–30851 (1997).
8. Presley, J. F. *et al.* *Nature* **417**, 187–193 (2002).
9. Yang, J. S. *et al.* *J. Cell Biol.* **159**, 69–78 (2002).
10. Weiss, M. & Nilsson, T. *Traffic* **4**, 65–73 (2003).
11. Desai, A. & Mitchison, T. J. *Annu. Rev. Cell Dev. Biol.* **13**, 83–117 (1997).
12. Baumgart, T., Hess, S. T. & Webb, W. W. *Nature* **425**, 821–824 (2003).

Chemistry

Internal combustion

Jeffrey Reimer

It's sometimes difficult to observe combustion *in situ* — inside, say, a porous material or an industrial reactor. But with the help of nuclear magnetic resonance, a new vista has opened up.

It was a burning candle that introduced generations of scientists to the discipline of observation. From the glow of the wick to the hues and flickers of the flame, combustion became the point of entry into the world of experimental science. Combustion continues to be a passion for many, who are drawn to the rich interplay of kinetics, thermodynamics and transport phenomena that describe modern combustion technologies.

Despite such allure, *in situ* experimental observations of combustion remain difficult in environments inaccessible to light. Measurements of reactions within porous media are particularly problematic. For example, many reactors are filled with solid particles to combust fuel catalytically (thereby reducing the combustion temperature and lowering emission of environmentally harmful nitrogen oxides); but the presence of these solids makes experimental access to temperature, pressure and composition inside the reactor very difficult. Looking to the future, nanoscale combustion engines embedded in a silicon chip will not be easily monitored using current combustion diagnostics.

In a contribution to the *Journal of the American Chemical Society*, Anala *et al.*¹ demonstrate the potential of nuclear magnetic resonance (NMR), one of the most effective toolkits in experimental science, to study combustion reactions and transport in opaque media. Magnetic resonance methods, both spectroscopic and imaging, rely on the collective observation of the angular momentum of around 10^{18} nuclei. Such huge numbers of nuclei are needed because

the energy differences between the various angular-momentum states of individual nuclei — which are the basis of the measurement — are small compared with the thermal energy available at room temperature. But this requirement hinders the acquisition of quantitative NMR data in gas-phase systems. To counteract this, Anala *et al.* have exploited a clever technique: xenon gas can be prepared so that individual atoms have angular momenta characteristic of extremely low temperatures, and then delivered to the NMR experiment at ambient temperatures (for a review, see ref. 2); using this 'hyperpolarized' xenon overcomes the sensitivity limitations inherent in such measurements.

Anala *et al.*¹ were thus able to detect NMR signals from xenon atoms in a methane flame burning inside a porous material (a zeolite molecular sieve). Their measurements reveal differing zones of temperature and pressure within the flame. For example, from the subtle shifts in the many resonance frequencies of the xenon atoms the authors determined the temperature changes experienced by xenon atoms inside the micropores of the material. They also noted slight changes in pressure experienced by xenon atoms in and above the bed of porous material.

In a second experiment, Anala *et al.* monitored the exchange of xenon atoms between different locations in the reaction region, manipulating and storing the observables associated with nuclear angular momenta using specially crafted radio-frequency pulses³ and time delays. In this scheme,

the time evolution of the xenon angular momenta is sensitive to the physical exchange of nuclei with different resonance frequencies. First, the researchers associate the different frequencies with positions inside the combustion apparatus; then they monitor the motion of xenon nuclei between different points in the flame. By increasing the time delay between NMR snapshots and thus the time allowed for the atoms to move, Anala *et al.* determined a timescale for the exchange of xenon atoms between the various environments of the flame, as shown in Fig. 1. After 5 milliseconds, for example, there is no atom exchange; after 20 milliseconds, xenon atoms inside the pores of the zeolitic material have been exchanged with other xenon atoms between the porous particles; and after 100 milliseconds, these in turn have also been exchanged with atoms above the bed of porous particles, yet still within the flame.

Although more research must be done to extend the strategies presented by Anala *et al.*¹, the potential of NMR techniques in the characterization of catalytic combustion is vast. For example, catalytic cracking of hydrocarbons in the petroleum industry

occurs in large industrial reactors that lack *in situ* characterization. 'Hot zones' — regions of extensive carburization — and other process gradients lead to inefficiencies in process control and in optimization of these reactors. Through NMR, these phenomena could be quantified. In addition to the use of hyperpolarized xenon, other techniques are gaining recognition — such as those using parahydrogen-generated hyperpolarization⁴ with chemical reactions to produce 'cold' nuclei in hydrocarbon molecules. A growing body of research is also providing more sensitive magnetometers^{5,6} for detecting nuclear angular momentum. These, and other new approaches, offer the exciting possibility of conducting NMR experiments on combustors that are not confined to the high-field magnets normally associated with magnetic resonance measurements^{7,8}.

The opportunity to use the methods described by Anala *et al.*¹ to study micro-engines — such as rotary-engine micro-machines fabricated on silicon substrates⁹ — should prove particularly interesting. For example, by using the hyperpolarized xenon as a fuel diluent and synchronizing the NMR experiment with the rotation of the engine parts (rather as clinicians do when

monitoring the human heart¹⁰), NMR information could be obtained that is sensitive to temperature and pressure within each mini-combustion chamber. Other forms of hyperpolarization may lead to the quantification and speciation of longer-lived reaction intermediates. Such data would be useful in the design of new engines that could one day replace the high-cost batteries in many consumer products. ■

Jeffrey Reimer is in the Department of Chemical Engineering, University of California, Berkeley, California 94720-9989, USA.

e-mail: reimer@socrates.berkeley.edu

1. Anala, S. *et al.* *J. Am. Chem. Soc.* **125**, 13298–13302 (2003).
2. Goodson, B. J. *Magn. Reson.* **155**, 157–216 (2002).
3. Ernst, R. R., Bodenhausen, G. & Wokaun, A. *Principles of Nuclear Magnetic Resonance in One and Two Dimensions* (Oxford Univ. Press, 1987).
4. Natterer, J. & Bargon, J. *Prog. Nucl. Magn. Reson. Spectrosc.* **31**, 293–315 (1997).
5. Augustine, M. P., TonThat, D. M. & Clarke, J. *Solid State Nucl. Magn. Reson.* **11**, 139–156 (1998).
6. Kominis, I. K., Kornack, T. W., Allred, J. C. & Romalis, M. V. *Nature* **422**, 596–599 (2003).
7. Blumich, B. *NMR Imaging of Materials* (Monogr. Phys. Chem. Mater. 57; Oxford Univ. Press, 2000).
8. Meriles, C. A., Sakellariou, D., Heise, H., Moulé, A. & Pines, A. *Science* **293**, 82–85 (2001).
9. Fernandez-Pello, A. C. *et al.* *Proc. Combustion Inst.* **29**, 883–899 (Reinhold, New York, 2003).
10. Kohler, S., Hiller, K. H., Jakob, P. M., Bauer, W. R. & Haase, A. *Magn. Reson. Med.* **50**, 449–452 (2003).

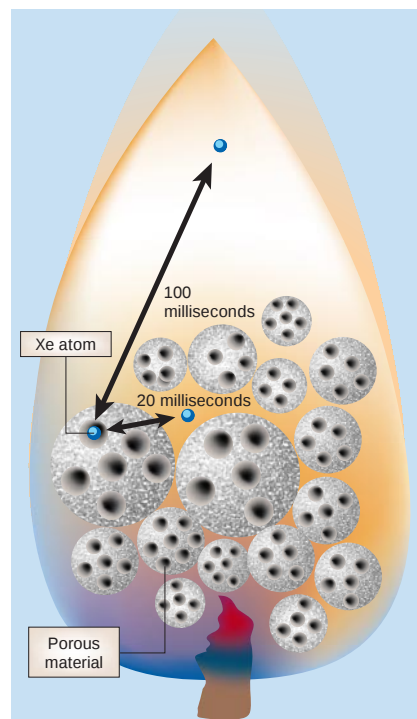


Figure 1 Flame proof. Using nuclear magnetic resonance, Anala *et al.*¹ have traced the motion of 'hyperpolarized' xenon atoms inside and around a bed of porous (zeolitic) pellets, as a methane flame burns within the material. On timescales of around 20 milliseconds, the xenon atoms are detected to have moved, swapping places with others inside the pores of the zeolite. After around 100 milliseconds, exchanges have occurred over longer distances, between the bed of material and the space above it.

Mammalian evolution

Isolationist tendencies

Jean-Jacques Jaeger

For some 40 million years, the Afro-Arabian landmass existed in splendid isolation. A newly described fossil fauna from the end of that time provides a window on the evolution of the continent's large mammals.

During most of the Cenozoic era, from the Cretaceous–Tertiary boundary 65 million years ago until roughly 24 million years ago, Afro-Arabia was an island continent drifting steadily northwards towards Eurasia¹. Fossil mammals documenting this period are scarce and belong almost exclusively to endemic forms restricted to Afro-Arabia, such as proboscideans², hyraxes³ and elephant-shrews. But by around 24 million years ago, a permanent land bridge had formed between the two landmasses⁴. A burst of faunal interchange followed: many Eurasian mammals, such as rhinos and ruminants, dispersed into Africa, and some Afro-Arabian mammals, such as elephants, migrated in the opposite direction.

On page 549 of this issue⁵, Kappelman *et al.* provide our first glimpse into the period just before this major biogeographical event. From precisely dated 27-million-year-old rocks in the Chilga region of the Ethiopian highlands, they describe the first known Afro-Arabian community of large herbi-

vores from the late Oligocene (Fig. 1, overleaf). This fauna clearly testifies to the continued isolation of Afro-Arabia at that time, for it includes endemic forms only — five distinct proboscideans, for example, three hyraxes and one arsiniothere (an extinct, rhino-like herbivore that was characterized by a pair of huge, divergent horns). It was known that several ancient lineages still survived around 27 million years ago. But the fossils from Chilga are especially valuable for the relative abundance of taxonomic groups and the first occurrence of modern representatives of some of the endemic mammals.

Among the proboscideans recorded are primitive forms such as *Palaeomastodon* and *Phiomia* (also known from older deposits in Egypt). But there are also representatives of modern families, for example taxa such as *Gomphotherium*, the earliest proboscidean on the branch leading to extant elephants. Another surprise is the oldest occurrence of deinotheres, peculiar proboscideans with downward-curved lower tusks, which were previously recorded only from rocks

younger than 24 million years old. The new species of deinother displays molars that are more 'bunodont' in form (that is, made of several distinct cusplets arranged in transverse crests) than its descendant, whose molars display plain transverse crests. This discovery seems to rule out the possibility that deinotheres are derived from an ancestor bearing plain, transverse-crested molars, as was formerly supposed, and provides new evidence about proboscidean evolution.

Most of our knowledge of Afro-Arabian mammalian communities in the earliest Tertiary has come from the Fayum desert in north-west Egypt. This region has yielded fossil-rich localities ranging in age from about 38 million to 32 million years ago^{6,7}. A few other sites, older in age, are known from north-west Africa⁸. All in all, however, places containing evidence of early African mammals are few and far between.

Nonetheless, considerable information has been inferred from the evidence we do have. First, several groups of extant mammals, including proboscideans, sirenians, hyraxes, aardvarks, tenrecs, elephant-shrews and golden-moles, originated on Afro-Arabia, indicating that there was a long period of endemic evolution. This conclusion has been supported by molecular data, leading systematists to unify these mammalian groups into the superorder Afrotheria (a name derived from the hypothesis that they all share a common ancestor and a long history in Africa⁹). Between 55 million and 24 million years ago, African mammalian faunas are dominated by these endemic forms. Only a few other mammals, such as rodents¹⁰ and primates¹¹, and a group known as anthracotheriid artiodactyls¹², which migrated during limited dispersals from adjacent continents, lived alongside these afrotheres. The rapid evolution of these newcomers led to early speciation on the Afro-Arabian landmass: the catarrhines¹³, a branch of primates ancestral to both hominoids and Old World monkeys, are one such case.

The Chilga mammals also yield insights into the dynamics of the faunal interchange between Afro-Arabia and Eurasia. As with the great American interchange, which occurred when the emergence of Central America in the late Tertiary brought North and South American mammals into contact¹⁴, the ensuing ecological competition ended with winners and losers. In this case, the peculiar *Arsinoitherium* appears to be documented for the last time. Hyraxes, which underwent a remarkable diversification³ in Afro-Arabia during the early Tertiary and occupied the ecological niches of other of the continent's herbivores, were greatly reduced in diversity. Proboscideans, by contrast, which had already started to undergo evolutionary radiation, continued to diversify, and the new land bridge offered them a unique opportunity to carry the

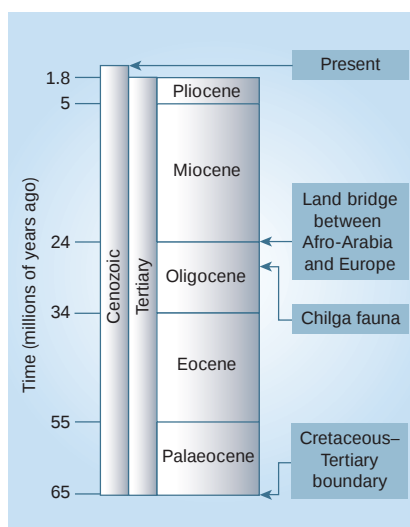


Figure 1 **Timeline: the Tertiary period, and date of the Chilga fauna, in context.**

process into Eurasia and, later, the Americas.

Finally, the discoveries of Kappelman *et al.*⁵ highlight two other palaeobiological issues. First, on northern continents glaciation caused a significant cooling around 33 million years ago, which resulted in numerous extinctions among mammalian communities¹⁵. From these new data, however, it seems that large Afro-Arabian herbivores were not affected, either at that time¹⁶ or later, implying that the climatic changes were less severe on southern continents. Second, the fossil record of the Afro-Arabian continent is not only scanty but also largely concentrated on the northern edge. This has led to the proposal that other groups of

mammals existed in Afro-Arabia during its period of isolation, but that they were restricted to more southern latitudes. However, the Chilga mammal community is rather like that found at Fayum in Egypt, which is some five million years older, providing hints that there was little provinciality among Afro-Arabian mammals at that time. As yet, though, we have unveiled only a few of the secrets of mammal evolution on the Afro-Arabian continent. Many more surprising discoveries are to be expected.

Jean-Jacques Jaeger is at the Institut des Sciences de l'Évolution, Université Montpellier II, CC64, Place Eugène Bataillon, 34095 Montpellier, France.

e-mail: jaeger@isem.univ-montp2.fr

- Smith, A. G., Smith, D. G. & Funnell, B. M. *Atlas of Mesozoic and Cenozoic Coastlines* (Cambridge Univ. Press, 1994).
- Shoshani, J. *et al.* in *The Proboscidea: Evolution and Palaeoecology of Elephants and Relatives* (eds Shoshani, J. & Tassy, P.) 57–75 (Oxford Univ. Press, 1996).
- Rasmussen, D. T. & Simons, E. L. *J. Vert. Paleontol.* **20**, 167–176 (2000).
- Jolivet, L. & Faccenna, C. *Tectonics* **19**, 1095–1106 (2000).
- Kappelman, J. *et al.* *Nature* **426**, 549–552 (2003).
- Simons, E. L. & Chatrath, P. S. in *Proc. Egyptian Geol. Surv. Centennial Conf. Spec. Publ. Geol. Surv. Egypt* **75**, 775–783 (1998).
- Gingerich, P. D. *J. Hum. Evol.* **24**, 207–218 (1993).
- Tabuce, R. Thesis, Univ. Montpellier II (2001).
- Stanhope, M. J. *et al.* *Mol. Phylogenet. Evol.* **9**, 501–508 (1998).
- Marivaux, L., Vianey-Liaud, M. & Jaeger, J.-J. *Zoologica Scripta* **31**, 225–239 (2002).
- Seiffert, E. R., Simons, E. L. & Attia, Y. *Nature* **422**, 421–424 (2003).
- Ducrocq, S. *Stuttgarter Beiträge zur Naturkunde B* **250**, 1–44 (1997).
- Seiffert, E. R. & Simons, E. L. *J. Hum. Evol.* **41**, 577–605 (2001).
- Lessa, E. P., Van Valkenburg, B. & Farina, R. A. *Palaeogeogr. Palaeoclim. Palaeoecol.* **135**, 1–4, 157–162 (1997).
- Prothero, D. R. *Annu. Rev. Earth Planet. Sci.* **22**, 145–165 (1994).
- Rasmussen, D. T., Bown, T. M. & Simons, E. L. in *Eocene–Oligocene Climatic and Biotic Evolution* (eds Prothero, D. R. & Berggren, W. A.) (Princeton Univ. Press, 1992).

Condensed-matter physics

Illuminating behaviour

Marc Bockrath

In fewer than three dimensions, the behaviour of electrons in metals should change to that of a 'Tomonaga–Luttinger liquid'. A photoemission study of one-dimensional carbon nanotubes supports this prediction.

Paraphrasing P. W. Anderson's famous phrase¹ that "more is different", many nanoscience researchers have come to the conclusion that "less is different", at least when it comes to the size of materials. Materials of greatly reduced dimension often behave in qualitatively different ways from their macroscopic counterparts. Much of this new behaviour arises in systems so small that the conduction electrons are no longer free to move in all three dimensions but are effectively confined to lower-dimensional spaces. In these low-dimensional systems, the effects of interactions between electrons can lead to fascinating new phenomena not observed in higher dimensions.

On page 540 of this issue, striking results are reported by Ishii *et al.*² from photoemission studies of carbon nanotubes, which are effectively one-dimensional systems. The data show a marked departure from the behaviour expected for non-interacting electrons, and provide a spectacular demonstration of the effects of interactions in low-dimensional systems.

Theoretical calculations indicate that, in a one-dimensional electron system such as a single-walled carbon nanotube (SWNT), the electrons exist in a state called a Tomonaga–Luttinger liquid (TLL)³. In a TLL, the low-energy excitations of its electrons are collective and sound-like, involving

the correlated motion of many electrons, rather than the single-electron-like excitations found in conventional three-dimensional metals. This has profound effects on many properties of the system. One such property is the 'single-particle spectral function', a quantity that is used to predict the results of experiments when an electron is suddenly added to or removed from an electron system. For a three-dimensional metal, the spectral function is expected to be nearly constant near the Fermi level (the surface of the electron sea, at zero temperature). In contrast, in a TLL, this addition or removal becomes difficult for electrons near the Fermi level, because it requires a complex rearrangement of all the other electrons in the system. As a result, the spectral function approaches zero at the Fermi level, following a power law with an exponent that depends on the strength of the interactions in the TLL. These interactions are conventionally parametrized by a dimensionless constant g , ranging between zero and one for repulsive interactions.

Numerous experiments to probe this behaviour have been performed on a variety of one-dimensional systems, from quantum Hall edge states⁴ to organic conductors⁵. SWNTs in particular have proved to be an excellent one-dimensional system, because electrons are able to travel large distances inside them without scattering^{6,7}. By attaching leads to the tubes, electrons can be injected into them, to probe the nanotube's single-particle spectral function. In such experiments, TLL behaviour should appear as a power-law dependence of the nanotube conductance on the applied voltage or temperature^{8,9}. Measurements on small bundles and single nanotubes have been in relatively good agreement with these predictions^{10,11}. However, transport measurements necessarily require at least two contacts, and electrons must travel along the nanotubes, potentially complicating the interpretation of the data.

The photoemission experiments of Ishii *et al.*² add significantly to the growing body of experimental literature on the subject. In these experiments, photons hitting the nanotubes liberate electrons from the sample. Energy conservation means that electrons that are more tightly bound emerge with less kinetic energy. So measuring the energies of emerging electrons yields a photoemission spectrum that shows how readily electrons at particular binding energies can be knocked out of the sample. As a result, photoemission is a direct probe of the single-particle spectral function.

Ishii *et al.* have used this technique to probe macroscopic quantities of nanotube material, consisting of bundles of 1.4-nm-diameter nanotubes. Each bundle contains a mixture of tubes that are either semiconducting or metallic, depending on the

precise diameter and chirality of the individual tubes. For electron binding energies above about 0.3 electronvolts in the photoemission spectrum, three peaks are observed that arise from the nanotube band structure¹²: the lower two binding-energy peaks are generated by semiconducting tubes, and the highest peak by metallic tubes. Careful analysis of the data indicates that about two-thirds of the nanotubes are semiconducting and the rest are metallic.

At low energies, considerations based on band structure predict a constant spectral function, arising from the metallic tubes. However, in sharp contrast, the experimental results show that the photoemission spectrum follows a power law in energy, extrapolating to zero at the Fermi level. The photoemission-spectrum intensity at the Fermi level follows similar behaviour, decreasing as a power law as the temperature is lowered, and extrapolating to zero at zero temperature. This is precisely what is expected qualitatively for a TLL. Quantitatively, the exponent of the power law can be predicted by calculation of the interaction parameter g in nanotubes. Theory predicts that g is 0.2, giving an expected exponent of about 0.4 for an isolated metallic SWNT^{8,9}. The experiments show excellent agreement with these calculations — the measured exponent is 0.48. Altogether, this provides an impressive demonstration of TLL behaviour in nanotubes.

It might seem surprising that the results from a heterogeneous mixture of tubes agree so well with theoretical predictions for a single nanotube. There are two factors favouring this agreement. One is that most

tubes are semiconducting — the semiconducting tubes are relatively poor conductors that effectively separate the metallic tubes from each other. The other is that the theoretically predicted exponents depend only very weakly on the distance to nearby conductors. So the exponent found for a metallic tube in a bundle should be similar to that of an isolated tube¹⁰.

These striking results represent only the tip of the large iceberg of fascinating properties that have been predicted for interacting electrons in one dimension. For example, such phenomena as spin transport¹³ and shot noise¹⁴ are expected to be drastically altered in TLL systems. These, and a wealth of other exotic properties of the TLL, await future exploration. ■

Marc Bockrath is in the Department of Applied Physics, California Institute of Technology, Pasadena, California 91125, USA.
e-mail: mwb@caltech.edu

1. Anderson, P. W. *Science* **177**, 393–396 (1972).
2. Ishii, H. *et al.* *Nature* **426**, 540–544 (2003).
3. Voit, J. *Rep. Prog. Phys.* **57**, 977–1116 (1995).
4. Chang, A. M., Pfeiffer, L. N. & West, K. W. *Phys. Rev. Lett.* **77**, 2538–2541 (1996).
5. Sekiyama, A., Fujimori, A., Aonuma, S., Sawa, H. & Kato, R. *Phys. Rev. B* **51**, 13899–13902 (1995).
6. Bachtold, A. *et al.* *Phys. Rev. Lett.* **84**, 6082–6085 (2000).
7. Liang, W. J. *et al.* *Nature* **411**, 665–669 (2001).
8. Kane, C., Balents, L. & Fisher, M. P. A. *Phys. Rev. Lett.* **79**, 5086–5089 (1997).
9. Egger, R. & Gogolin, A. O. *Eur. Phys. J. B* **3**, 281–300 (1998).
10. Bockrath, M. *et al.* *Nature* **397**, 598–601 (1999).
11. Yao, Z., Postma, H. W. C., Balents, L. & Dekker, C. *Nature* **402**, 273–276 (1999).
12. Dresselhaus, M. S., Dresselhaus, G. & Eklund, P. C. *Science of Fullerenes and Carbon Nanotubes* (Academic, New York, 1996).
13. Balents, L. & Egger, R. *Phys. Rev. B* **64**, 035310 (2001).
14. Bena, C., Vishveshwara, S., Balents, L. & Fisher, M. P. A. *J. Statist. Phys.* **103**, 429–440 (2001).

Medicine

Taking apart a cancer protein

Pier Paolo Scaglioni and Pier Paolo Pandolfi

Generation of a particular 'fusion' protein is characteristic of one type of leukaemia. But is it in fact the cleavage of this protein into smaller parts that is important? Provocative new findings suggest that it is.

Cancers develop when genetic mistakes accumulate, causing cells to proliferate unchecked and to show a reluctance to die when required. In many human leukaemias, one of the initial genetic errors is a chromosomal translocation: one part of a chromosome fuses with another, creating a composite gene with unusual properties. A well-known example of this occurs in acute promyelocytic leukaemia (APL), a cancer in which promyelocytes — a type of bone-marrow cell that produces certain white blood cells — accumulate in the bone marrow and blood. In most patients with APL, the cancerous cells harbour a chromosome translocation that fuses the

promyelocytic leukaemia (PML) gene to the gene for retinoic acid receptor- α (RAR α), giving rise to two fusion proteins: PML-RAR α and RAR α -PML¹. It had been thought that full-length PML-RAR α is required for this cancer to develop. Writing in *Cell*, however, Lane and Ley² challenge this concept. They find that this fusion protein is cleaved at several positions by neutrophil elastase — and that eliminating this enzyme in mice offers some protection from APL.

It is generally believed that the chromosomal translocations characteristic of APL — and those in other types of leukaemia — occur in immature blood-cell precursors that can self-renew³. In this way, the fusion

proteins produced by the translocations would lead to the expansion of a pool of cancer cells. Animal models support this idea: for instance, expressing PML-RAR α in immature myeloid precursors in mice leads to the expansion of these cells, blocks their differentiation programme at the promyelocytic stage, and results in leukaemia. Expressing the fusion protein in mature myeloid cells, however, does not result in leukaemia. These observations hint that the production of PML-RAR α is a key event in the development of APL, and that its activity is restricted to just a subset of immature myeloid cells⁴ (although it can also cause cancer when expressed in other tissues, including the skin and liver^{5,6}).

There is much evidence to support the notion that the full-length form of this fusion protein is necessary to cause APL. But Lane and Ley² propose another idea. They have found that the protein is sliced into several pieces by neutrophil elastase, an enzyme that is expressed most strongly in promyelocytes. The authors provide evidence that such cleavage occurs in the nucleus of mouse bone-marrow-derived cells and, more importantly, in cells taken from human and mouse APL. They identified short cleavage products in the lysates of APL samples (although, surprisingly, not in NB4 cells, which are derived from APL).

Lane and Ley also studied the incidence of APL in mice lacking neutrophil elastase. They crossed these mice with animals that had been genetically engineered to develop APL after a long latency period. Strikingly, the offspring were protected, at least partially, from APL: only about half of them developed the cancer. The results suggest that neutrophil elastase promotes the development of this leukaemia, and that PML-RAR α may need to be cleaved to unleash its full potential as a cancer-causing agent.

These findings are provocative, for they stand at odds with decades of work showing that the full-length fusion protein promotes APL. Research from several laboratories has found that it does so by preventing PML and RAR α from carrying out their normal functions^{1,7}. Usually, RAR α , on binding to retinoic acid, activates several target genes that enable myeloid cells (such as promyelocytes) to differentiate into granulocytes (a type of white blood cell)⁸. Full-length PML-RAR α can still bind to both DNA and retinoic acid, but no longer activates gene expression; instead, it recruits various repressors through its PML portion. This aberrant protein effectively outcompetes the normal RAR α .

PML, on the other hand, normally accumulates in speckled subnuclear structures termed PML nuclear bodies, and inhibits tumour development in several ways⁹. The PML-RAR α fusion protein can interact physically with PML, disrupting the nuclear

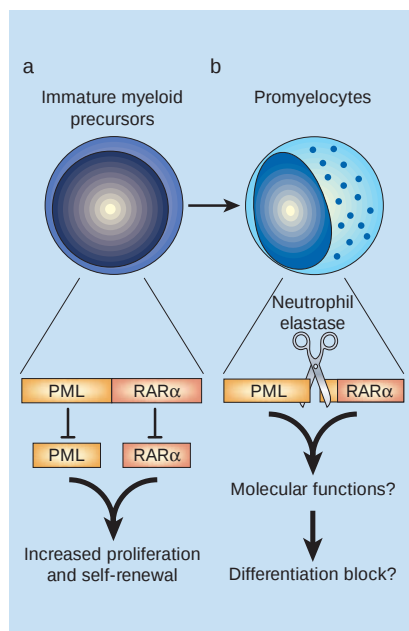


Figure 1 Hypothetical role of the PML-RAR α fusion protein in acute promyelocytic leukaemia (APL). The fusion protein is thought to contribute to this cancer by interfering with the normal functions of the individual proteins, PML and RAR α . Lane and Ley² have found, however, that cleavage of the fusion protein by neutrophil elastase in promyelocytes promotes the development of APL. One way of reconciling these findings could be to propose that a, the full-length fusion protein, when expressed in immature precursor cells, leads to increased proliferation and self-renewal. b, Cleavage of PML-RAR α might then generate products that block the differentiation of the larger pool of promyelocytes.

bodies. There is evidence to suggest that both of these effects of the fusion protein — on PML and on RAR α — are needed for leukaemia to develop (see, for example, ref. 9). So how do Lane and Ley's findings² fit in? If the cleavage of PML-RAR α by neutrophil elastase indeed plays a major part in causing APL, as the authors propose, then presumably the cleavage products also must block the normal functions of RAR α and PML. Whether or not they do remains to be seen. Alternatively, the cleavage products could have completely different functions that are yet to be uncovered.

The importance of the full-length fusion protein is also underscored by the fact that two effective drugs to treat this cancer, retinoic acid and arsenic trioxide, work by promoting the wholesale degradation of PML-RAR α . This has the effect of relieving the repression of RAR α 's target genes — so promoting differentiation — and allowing the PML nuclear bodies to re-form. If the products of PML-RAR α cleavage indeed contribute to leukaemia, these drugs must presumably also work against the cleavage

products. Do they? Again, that's a question for the future.

Another issue concerns the fact that neutrophil elastase seems to be most active in promyelocytes. But the cells that maintain the pool of cancerous promyelocytes are more likely to be less mature precursors (because these are the cells that proliferate) — and presumably neutrophil elastase is less active in these cells. On the other hand, a lack of this enzyme evidently does protect mice from the leukaemia-promoting effects of PML-RAR α . One way of reconciling these findings would be to propose that the full-length product gives the immature myeloid precursors increased proliferative and self-renewal potential¹⁰, and that the cleavage products act on the later promyelocytes, exacerbating the differentiation block (Fig. 1). This would also explain why PML-RAR α promotes cancer in cells that lack neutrophil elastase activity (such as liver cells).

A number of future lines of investigation suggest themselves. First, we need to characterize the molecular and biological properties of the major and minor products of PML-RAR α cleavage. Sequence analysis of the two major products hints that they may still interact physically with normal PML and be able to bind DNA through their RAR α portion; if so, they might still interfere with the functions of PML and RAR α . Expressing the cleavage products, or neutrophil-elastase-resistant PML-RAR α , in mice or in bone-marrow stem cells should provide some answers. Second, we need to confirm the existence of elastase-generated PML-RAR α products in a large cohort of patients. Finally, it will be interesting to see whether the pattern of cleavage products is affected by treatment drugs, and whether a specific pattern might predict clinical outcome and response to therapy. Lane and Ley's provocative findings and their ramifications look set to keep the APL scholars busy for years to come. What the implications are for cancer patients remain to be seen. ■

Pier Paolo Scaglioni and Pier Paolo Pandolfi are in the Molecular Biology Program, Departments of Medicine and of Pathology, Sloan-Kettering Institute, Memorial Sloan-Kettering Cancer Center, New York, New York 10021, USA.
e-mail: p-pandolfi@ski.mskcc.org

1. Melnick, A. & Licht, J. D. *Blood* **93**, 3167–3215 (1999).
2. Lane, A. A. & Ley, T. J. *Cell* **115**, 305–318 (2003).
3. Reya, T., Morrison, S. J., Clarke, M. F. & Weissman, I. L. *Nature* **414**, 105–111 (2001).
4. Piazza, F., Gurrieri, C. & Pandolfi, P. P. *Oncogene* **20**, 7216–7222 (2001).
5. Hansen, L. A. et al. *Cancer Res.* **63**, 5257–5265 (2003).
6. David, G., Terris, B., Marchio, A., Lavau, C. & Dejean, A. *Oncogene* **14**, 1547–1554 (1997).
7. Lin, R. J., Sterns, T., Tini, M. & Evans R. M. *Oncogene* **20**, 7204–7215 (2001).
8. Kastner, P. & Chan, S. *Oncogene* **20**, 7178–7185 (2001).
9. Salomoni, P. & Pandolfi, P. P. *Cell* **108**, 165–170 (2002).
10. Minucci, S. et al. *Blood* **100**, 2989–2995 (2002).

Cancer

Germ-cell suspect in the male

Cancer Cell **4**, 361–370 (2003)

Testicular tumours, the most common form of cancer in young men, usually develop when the reproductive germ cells begin dividing uncontrollably. The cellular events that trigger the process are largely unknown.

Sharon Gidekel and colleagues have implicated a protein, called Oct-3/4 and known to help regulate sperm production, in tumour formation. Gidekel *et al.* found that it was present in all human testicular germ-cell tumours that they tested, and occurred even in early-stage tumours that had not yet become malignant.

Protein levels influenced malignancy in a dose-dependent manner. Mice injected with cells expressing high levels of Oct-3/4 developed highly malignant tumours; those receiving cells with Oct-3/4 at low levels developed less aggressive tumours.

When Oct-3/4 concentrations were directly reduced in mouse tumours, the growths went into remission. Gene-directed therapies that eliminate the protein may prove useful against human testicular cancers, the authors speculate. In adults, Oct-3/4 is expressed only in germ cells, so Gidekel *et al.* suspect that treatments that target it directly should not affect other tissue types.

Helen R. Pilcher

Optics

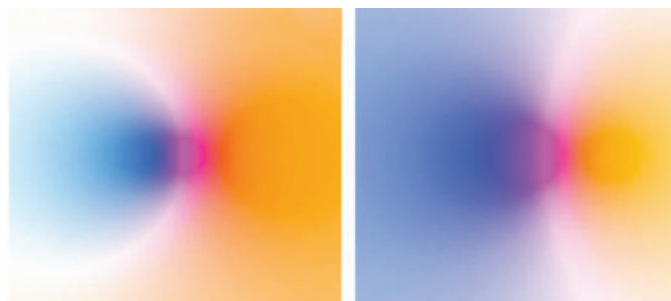
Colourful 'black holes'

New J. Phys. **5**, 154.1–154.7 (2003)

What is the colour of darkness? Proving that this is not some Zen koan, Jonathan Leach and Miles J. Padgett have teased out the chromatic patterns around an 'optical vortex', where a white-light beam has zero intensity and indefinite phase.

Such vortices are examples of phase singularities, 'dark' spots caused by wave interference. They also occur, for example, in ocean tidal patterns. In 2002, Michael Berry predicted that singularities in white light are surrounded by distinctive patterns of colour: a blue and an orange-yellow region separated by a purplish circular boundary, with green notably absent (see picture). Leach and Padgett have now confirmed that this chromatic structure appears around optical vortices, created by passing light from a tungsten halogen bulb through a liquid-crystal diffraction grating with a Y-shaped slit pattern.

Because of the very low intensity close to the singularity, it is no simple matter to convert the observed colour pattern to a chromatic image that corresponds to what a sufficiently sensitive human eye would



Out of the darkness: white light, here from black-body radiators at 3,300 K (left) and 6,000 K (right), is still colourful around 'dark' vortices.

register: the researchers used a CCD camera with a trichromatic response matched to the colour response of the eye.

Philip Ball

Climate

Warmth, wind and water

J. Clim. **16**, 3793–3802 (2003)

Why is there an evaporation minimum at the Equator? Richard Seager and colleagues ask themselves this question, and tackle it with computer models of atmospheric and oceanic behaviour.

At all points on the Equator there is less evaporation than in neighbouring tropical regions both north and south. In some places this can be largely explained by the consequences of a cold tongue of upwelling water. But this mechanism cannot apply to the Indo-Pacific warm pool, the site of the warmest waters in the world.

Proposed explanations are put in terms of exchanges of latent heat, and two in particular are accepted. One is that cloud cover over the warm pool reduces the amount of solar radiation reaching the surface, thereby reducing the balancing flux of latent heat required. The other is that less vigorous winds, compared with the trade winds to the north and south, have the same effect.

Seager and colleagues' modelling work brings in another consideration: they find that the wind factor can be accounted for only when heat transport in the ocean itself is included in the picture, through divergence of the heat flux away from the warm pool. But the authors candidly admit that the extent of atmosphere–ocean coupling requires further study.

Tim Lincoln

Ecology

Eaten history

Mol. Ecol. **12**, 3467–3475 (2003)

In arable fields, spiders help in pest control by consuming large numbers of aphids. But aphids have poor nutritional value, and are sometimes toxic. By investigating the contents of spiders' stomachs using DNA-based techniques, N. Agustí and colleagues provide a new angle on the importance of springtails (Collembola) in the spider diet.

The researchers developed genetic markers that allowed them to amplify and identify the mitochondrial DNA of three species of Collembola in the gut contents of money spiders. They say that this is the first time such techniques have been used to identify

prey species from the guts of wild-living arthropods.

The spiders preferred to eat the largest collembolan, *Isotoma anglicana* — this springtail was heavily represented in their stomachs, even though it is relatively rare in the wild. Agustí and colleagues point out that farming methods that increase the population density of this species could boost spider numbers, and so have a beneficial knock-on effect on pest control.

John Whitfield

Developmental biology

Serpin takes its Toll

Dev. Cell doi:10.1016/S1534580703003381 (2003)

Curr. Biol. doi:10.1016/S0960982203008303 (2003)

The back of most creatures, the dorsal side, is different to the belly, or ventral side. In the fruitfly *Drosophila*, such lopsided development stems in part from local activation of the Toll signalling pathway. Two groups — Carl Hashimoto *et al.* and Petros Ligoxygakis *et al.* — have now identified a serine protease inhibitor, a serpin, that is essential to blunt Toll signalling and confine it to ventral areas.

The Toll pathway is turned on by a cascade of proteases, enzymes that cut up other proteins. The fluid surrounding the *Drosophila* embryo contains proteases with names such as Gastrulation defective, Snake and Easter, which cut and activate one another consecutively at the ventral side of the embryo. Finally, active Easter slices the protein Spätzle and a piece of the latter activates the Toll receptor protein. But activating Easter locally is not enough to confine Toll signalling ventrally.

Activation of the Toll pathway bears a striking resemblance to the blood-clotting cascade, in which proteases are turned on at the site of tissue damage and serpins inactivate the ones that diffuse away. Both groups followed up this connection in *Drosophila*. A serpin-like gene indeed emerged that seems necessary and sufficient to keep Easter activity in check in developing flies. These findings suggest that inhibition of proteases by serpins may be a general mechanism for spatially controlling biological processes.

Marie-Thérèse Heemels

On-chip manipulation of free droplets

Tiny free-floating drops can be driven across a liquid medium by an electric field.

Lab-on-a-chip' systems resemble factories with permanently rigged pipes, but their prefabricated microchannels could have problems in delivering materials such as suspended particles, biological cells or proteins, which may adhere to the walls and clog the channels. More flexible microfluidic systems allow liquids to be transported as droplets on a solid surface^{1–10}, but these suffer from similar drawbacks where the droplets are in contact with solid walls. Here we describe a liquid–liquid microfluidic system for manipulating freely suspended microlitre- and nanolitre-sized droplets of water or hydrocarbon, which float on a denser, perfluorinated oil and are driven by an alternating or constant electric field applied by arrays of electrodes below the oil. These microfluidic chips could be used as a versatile tool in microscale transport and mixing and in chemical and materials synthesis.

In our liquid–liquid microfluidic device (Fig. 1), water or dodecane microdroplets float freely on a surface of fluorinated oil ('F-oil') and alternating (a.c.) and/or constant (d.c.) electric fields are applied through electrodes below the surface of the oil phase. Spatially inhomogeneous a.c. fields evoke dielectrophoretic force¹¹, which attracts polarizable objects to areas of high field intensity.

The droplets migrate to the energized electrodes and hover above them, trapped by the field (Fig. 2a). These droplets are readily moved by programmed switching of the voltage to different electrodes (see Fig. 2a, b and supplementary information). The voltages applied (200–600 V) are comparable to those used in other dielectrophoretic experiments. The droplet speed

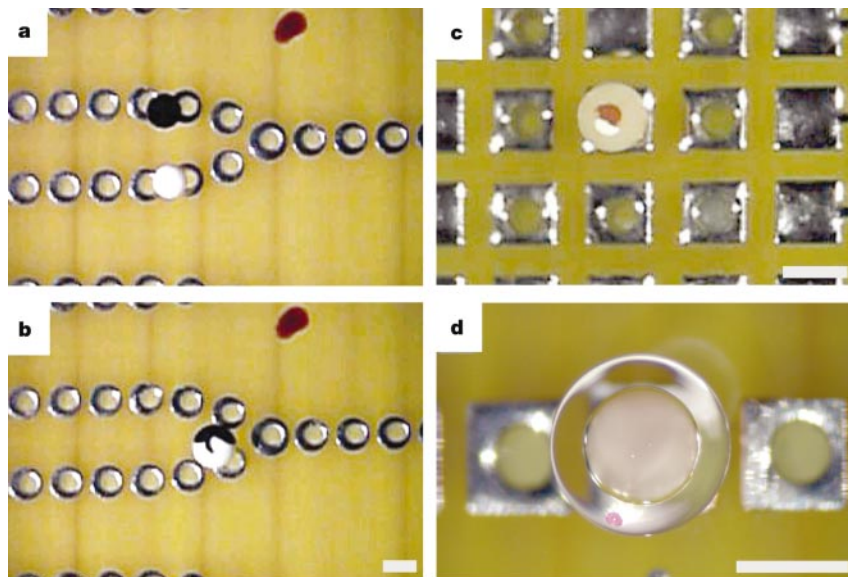


Figure 2 Behaviour of freely suspended droplets above a liquid–liquid microfluidic chip (for movies, see supplementary information). **a**, Droplets (750 nanolitres) containing latex microspheres (white) and gold nanoparticles (purple) are transported; **b**, the droplets mix at the moment they touch near the track junction. The mixed droplet then continues further along the single track. **c**, Two droplets containing polystyrene (white) and magnetic (brown) latexes temporarily form an anisotropic-particle aggregate on mixing. **d**, Water droplet containing latex and 2 mM sodium dodecyl sulphate is encapsulated inside a liquid dodecane shell. Scale bars, 1 mm. Further details are available at <http://crystal.che.ncsu.edu>.

in a.c. mode scales with the square of the field intensity, in agreement with dielectrophoretic theory¹¹. Because the droplets encounter only viscous resistance, they move with extremely low power dissipation, albeit more slowly than when in contact with electrodes on solid walls^{8–10}.

Water droplets are also attracted or repelled strongly by d.c. fields, moving with velocities as high as 2.0 mm s^{−1}. This indicates that they have significant charge and/or dipole moments and move by

coulombic repulsion or attraction. Droplet charging and recharging, which possibly occurs by ion transfer through the F-oil, were observed routinely. We also found evidence of unexpectedly strong internal polarization in the droplets (to be discussed elsewhere). The strong charging effects allow hydrocarbon oil droplets to be manipulated as well; however, these do not respond to symmetric a.c. fields because of their low polarizability.

Crystalline shell-like balls form by precipitation of inorganic solids inside mixed droplets, and asymmetric microassemblies^{9,12} result from drying of droplets containing microparticles and nanoparticles (Fig. 2c; see supplementary information); the chips could be used to manipulate the resulting suspended solid particles as well. By combining dodecane and water droplets in the presence of surfactant, we were able to encapsulate the water symmetrically inside a hydrocarbon shell (Fig. 2d, and see supplementary information).

The flexible microfluidic devices described here would be useful in a range of applications — for example, in the synthesis of materials and in biological microassays. More complex applications might include confining living cells or genetic material into individual droplet containers, and performing biochemical

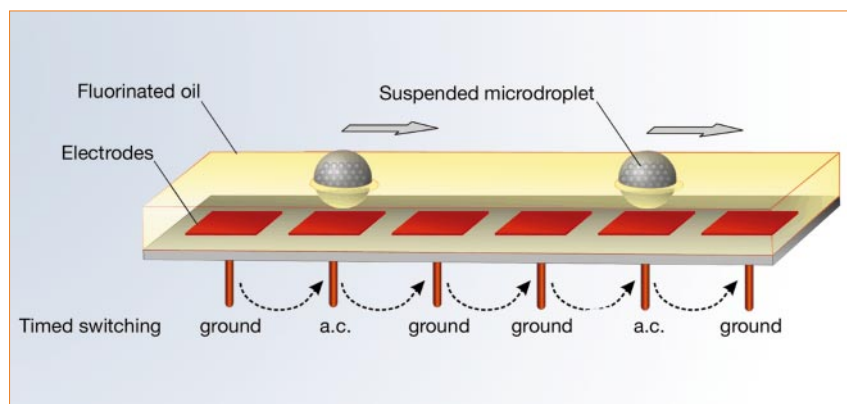


Figure 1 Schematics of the dielectrophoretic transporter used for the operation of a liquid–liquid microfluidic chip with freely suspended droplets. Water or dodecane droplets of volume 500–1,000 nanolitres float freely on the surface of perfluoromethyldecalin ('F-oil', an inert, dense liquid)¹³. Alternating (a.c.) and/or constant electric fields are applied by arrays of electrodes immersed in the F-oil phase a few millimetres below the surface.

reactions, precipitation assays or parallel drug or toxin screening.

Orlin D. Velev, Brian G. Prevo, Ketan H. Bhatt

Department of Chemical Engineering, North Carolina State University, Raleigh, North Carolina 27695-7905, USA

e-mail: odvelev@unity.ncsu.edu

1. Chaudhury, M. K. & Whitesides, G. W. *Science* **256**, 1539–1541 (1992).
2. Lee, S. W. & Laibinis, P. E. *J. Am. Chem. Soc.* **122**, 5395–5396 (2000).
3. Daniel, S., Chaudhury, M. K. & Chen, J. C. *Science* **291**, 633–636 (2001).
4. Ichimura, K., Oh, S. K. & Nakagawa, M. *Science* **288**, 1624–1626 (2000).
5. Daniel, S. & Chaudhury, M. K. *Langmuir* **18**, 3404–3407 (2002).
6. Washizu, M. *IEEE T. Ind. Appl.* **34**, 732–737 (1998).
7. Rosslee, C. & Abbott, N. L. *Curr. Opin. Coll. Interface Sci.* **5**, 81–87 (2000).
8. Pollack, M. G., Fair, R. B. & Shenderov, A. D. *Appl. Phys. Lett.* **77**, 1725–1726 (2000).
9. Jones, T. B., Gunji, M., Washizu, M. & Feldman, M. J. *J. Appl. Phys.* **89**, 1441–1448 (2001).
10. Cho, S. K., Moon, H. J. & Kim, C. J. *J. Microelectromech. S* **12**, 70–80 (2003).
11. Fuhr, G. et al. *Naturwissenschaften* **81**, 528–535 (1994).
12. Dinsmore, A. D. et al. *Science* **298**, 1006–1009 (2002).
13. Velev, O. D., Lenhoff, A. M. & Kaler, E. W. *Science* **287**, 2240–2243 (2000).

Supplementary information accompanies this communication on Nature's website.

Competing financial interests: declared none.

Palaeontology

A polydactylous amniote from the Triassic period

The earliest four-limbed vertebrates, or tetrapods, lived between 370 million and 354 million years ago, during the Late Devonian period, and typically had more than five digits (polydactyly)¹. We have discovered that a preaxial form of polydactyly, in which extra digits are positioned anterior to the first digit, has unexpectedly re-emerged in a marine reptile

from the Early Triassic period about 242 million years ago — the overall morphology of both the manus and pes closely resemble those of the earliest tetrapods. Until now, no post-Devonian tetrapod has been found with a comparative type of polydactyly, so the new amniote provides a striking example of convergent evolution.

The new amniote is a marine reptile from Hubei Province in China (amniotes include reptiles, birds and mammals). Its most remarkable feature is the number of its digits: the forelimbs bear seven and the hindlimbs have six (Fig. 1). The extra digits on both fore- and hindlimbs are well developed and the bones are arranged normally as distal carpal/tarsal (distal carpals 3 and 4 are fused as a single large carpal in the forelimb of the new amniote), metacarpal/metatarsal and phalanges (see supplementary information).

Ichthyosaurs of the Mesozoic era (250–65 million years ago) had porpoise-like bodies, with dorsal and tail fins and often polydactylous limbs. This polydactyly, however, was quite different from that of the Devonian tetrapods. Modern pandas and moles², humans^{2,3} and cats⁴ occasionally have extra preaxial digits, but these are rarely morphologically or structurally comparable with a normal digit^{2,3,5}.

Almost all polydactyly in tetrapods can be referred to one of three types. In postaxial polydactyly, the extra digits are posterior to digit V, as seen in the Late Devonian *Tulerpeton*⁶, some frog individuals and even humans^{2,3}. In preaxial polydactyly, the extra digits are anterior to digit I, as seen in the fore- and hindlimbs of the new amniote, and in the hindlimbs of the Late Devonian *Ichthyostega* and forelimbs of the Late Devonian *Acanthostega*¹. In bilateral polydactyly, the extra digits are anterior to

digit I and posterior to digit V, as seen in the forelimbs of the ophthalmosaurian ichthyosaurs⁷ and today in some polydactylous Indian families³; the extra digits on the hindlimbs of the Devonian *Acanthostega*⁸ also appear to be of this type.

Other types of polydactyly can occur, for example in the forelimbs of many non-ophthalmosaurian ichthyosaurs⁷. Most occur by interdigital or postaxial phalangeal bifurcation⁷. Of the known polydactylous tetrapods, the new amniote is the only one that has both fore- and hindlimbs that are preaxially polydactylous, matching the current limb-development model⁹ (see supplementary information).

The new amniote was a secondarily aquatic reptile and its polydactylous limbs are derived from adaptation to its aquatic life. Its manus and pes are short and wide, and generally resemble those of the Late Devonian *Ichthyostega* and *Acanthostega*. They are also comparable in shape to the limb-like paired fins of extant frogfishes^{8,10}. The limbs of this amniote may have functioned in a similar way to those of the Devonian tetrapods or to the paired fins of frogfishes when moving across underwater substrates. In its morphology and way of life, the new amniote provides a good example of evolutionary convergence with the earliest tetrapods, as well as an analogy with frogfishes in vertebrate evolution.

Xiao-Chun Wu*, **Zhan Li†**, **Bao-Chun Zhou‡**, **Zhi-Ming Dong‡**

*Canadian Museum of Nature, PO Box 3443, Station D, Ottawa, Ontario K1P 6P4, Canada e-mail: xcwu@mus-nature.ca

†Shanghai Science and Technology Museum, 2000 Century Avenue, Pudong District, Shanghai 200127, China

‡Institute of Vertebrate Paleontology and Paleoanthropology, Chinese Academy of Sciences, PO Box 643, Beijing 100044, China

1. Coates, M. I. & Clack, J. A. *Nature* **347**, 66–69 (1990).
2. Gould, S. J. *Nat. Hist.* **1/91**, 22–29 (1991).
3. Radhakrishna, U. et al. *Am. J. Genet.* **65**, 645–655 (1999).
4. Danforth, C. H. *Am. J. Anat.* **80**, 143–171 (1947).
5. Shubin, N., Tabin, C. & Carroll, S. *Nature* **388**, 639–648 (1997).
6. Lebedev, O. A. *Doklady Akademii Nauk SSSR* **278**, 1470–473 (1984).
7. Motani, R. *J. Vert. Paleontol.* **19**, 28–41 (1999).
8. Clack, J. *Gaining Ground: The Origin and Evolution of Tetrapods* (Indiana Univ. Press, Bloomington, 2002).
9. Shubin, N. H. & Alberch, P. *Evol. Biol.* **20**, 319–387 (1986).
10. Edwards, J. J. *Am. Zool.* **29**, 235–254 (1989).
11. Li, J.-L., Liu, J., Li, C. & Huang, Z.-X. *Vert. Palasiat.* **40**, 241–244 (2002).

Supplementary information accompanies this communication on Nature's website.

Competing financial interests: declared none.

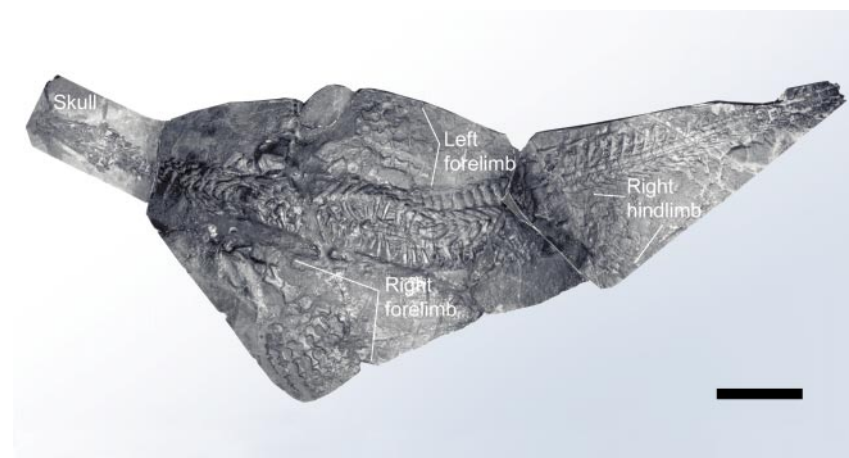


Figure 1 The Triassic polydactylous amniote, housed at the Shanghai Science and Technology Museum (specimen SSTM 5025). It is represented by the part and counterpart of an almost complete skeleton from which the anterior end of the snout and the tail tip are missing. The specimen was collected from the marine Jialingjiang Formation (late Early Triassic¹¹) near Xunjiang, Nanzhang County, Hubei Province, China. Taxonomically, SSTM 5025 is referred to Nanchangosauridae, Wang, 1959 of Hupehsuchia Young & Dong, 1972 in Diapsida Gauthier et al., 1988 of Reptilia Gauthier et al., 1988. Further details are available from X.-C. W. Scale bar, 10 cm.

brieF communications is intended to provide a forum for brief, topical reports of general scientific interest and for technical discussion of recently published material of particular interest to non-specialist readers (**communications arising**). Priority will be given to contributions that have fewer than 500 words, 10 references and only one figure. Detailed guidelines are available on Nature's website (www.nature.com/nature).

TRP channels as cellular sensors

David E. Clapham

Howard Hughes Medical Institute, Pediatric Cardiology, Children's Hospital of Boston, Department of Neurobiology, Harvard Medical School, Enders 1309, 320 Longwood Avenue, Children's Hospital, Boston, Massachusetts 02115, USA

TRP channels are the vanguard of our sensory systems, responding to temperature, touch, pain, osmolarity, pheromones, taste and other stimuli. But their role is much broader than classical sensory transduction. They are an ancient sensory apparatus for the cell, not just the multicellular organism, and they have been adapted to respond to all manner of stimuli, from both within and outside the cell.

The human genome encodes hundreds of channels that broker the passage of charged ions across impermeable lipid bilayers¹. While energy-requiring pumps labour to build charge and concentration gradients across the membrane, ion channels spend this stored energy, much as a switch releases the electrical energy of a battery. Small conformational changes cause channels to open, allowing over ten million ions to flow per second through each channel. Ca^{2+} ions are particularly important in cellular homeostasis and activity, and the surface of each cell holds thousands of channels that precisely control the timing and entry of Ca^{2+} ions.

Transient receptor potential (TRP) channels were first described in *Drosophila*, where photoreceptors carrying *trp* gene mutations exhibited a transient voltage response to continuous light^{2,3}. Unlike most ion channels, TRP channels are identified by their homology rather than by ligand function or selectivity, because their functions are disparate and often unknown. They have been called store-operated channels (SOCs), but this description is theoretical and related to a poorly understood phenomenon.

The known functions are diverse. Yeast use a TRP channel to

perceive and respond to hypertonicity^{4,5}. Nematodes use TRP channels at the tips of neuronal dendrites in their 'noses' to detect and avoid noxious chemicals⁶. Male mice use a pheromone-sensing TRP channel to tell males from females⁷. Humans use TRP channels to appreciate sweet, bitter and umami (amino acid) tastes⁸, and to discriminate warmth, heat and cold. In each of these cases, TRPs mediate sensory transduction, not only in a classical sense, for the entire multicellular organism, but also at the level of single cells. Almost all mammalian TRP channel genes are now known. Here I summarize the common characteristics of the diverse mechanisms of TRP channel activation, highlighting major questions that remain to be answered. More details can be found in other reviews^{9–15}.

What are TRP channels?

Mammalian TRP channels comprise six related protein families with sequence identity as low as 20% (Fig. 1). All TRP channels are putative six-transmembrane (6TM) polypeptide subunits that assemble as tetramers to form cation-permeable pores (Box 1). In general, they are almost ubiquitously expressed and most have splice variants. So most cells have a number of TRP channel proteins.

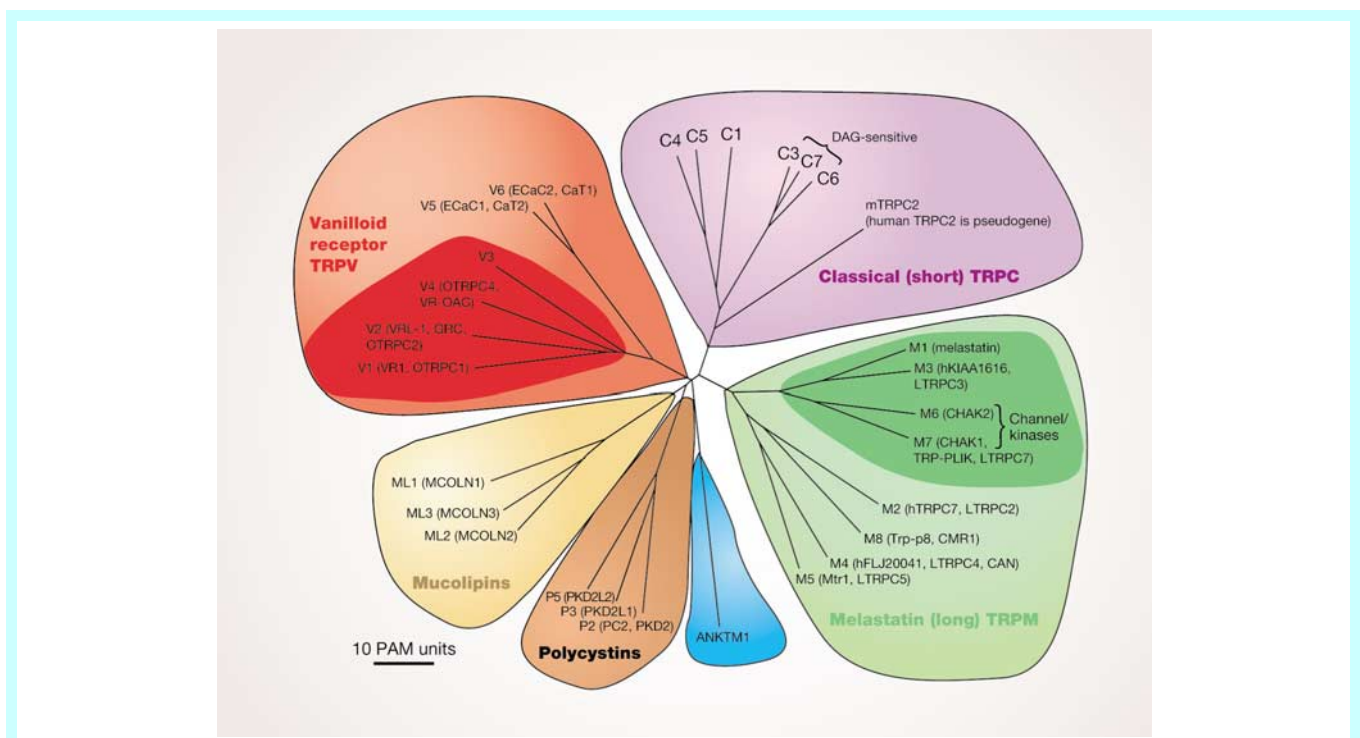


Figure 1 Mammalian TRP family tree. The evolutionary distance is shown by the total branch lengths in point accepted mutations (PAM) units, which is the mean number of substitutions per 100 residues.

It will take time to elucidate all of the diverse functions of TRP proteins. To date, the most informative approaches have been ligand-specific expression cloning, and targeted (global) gene inactivation in mice. In the near future, further information is likely to be gleaned from tissue-specific and developmental-stage-specific gene targeting.

The TRPC (canonical TRP) subfamily

All mammalian TRPC proteins appear to be analogous to the TRP involved in *Drosophila* phototransduction, in that they function as receptor-operated channels. They are activated by stimulation of G-protein-coupled receptors (GPCRs) and receptor tyrosine kinases (Box 2).

TRPC1, the first mammalian TRP reported¹⁶, forms heteromeric channels with TRPC4 and/or TRPC5 (ref. 10). The properties of the heteromultimers are distinct from those of TRPC4 and TRPC5 homomultimers (see Table 1 and Fig. 2). TRPC5, but not TRPC1, is present in hippocampal growth cones and modulates neurite

extension¹⁷. Mice lacking TRPC4 have defects in agonist-induced vasoregulation and lung microvascular permeability^{18,19}.

TRPC3, TRPC6 and TRPC7 proteins share ~75% identity, have relatively low selectivity for Ca^{2+} over Na^{+} , and are sensitive to the intracellular concentration of Ca^{2+} ($[\text{Ca}^{2+}]_i$). Diacylglycerol (DAG) analogues²⁰ potentiate their activity, but not through protein kinase C activation. TRPC3 has been investigated extensively as a putative inositol-1,4,5-trisphosphate (InsP_3) receptor (IP_3R)-binding SOC, with conflicting results²¹. All of the TRPC3, TRPC6 and TRPC7 subfamily are highly expressed in smooth and cardiac muscle cells, making them candidates for the receptor-activated nonselective cation channels known to exist in these sites. In support of this idea, TRPC6 is an essential part of the α_1 -adrenoreceptor-activated cation channel in rabbit portal vein myocytes²². They may also have roles in the regulation of vascular tone, airway resistance and cardiac function.

TRPC2 appears to be a pseudogene in humans, but its rat

Box 1

TRP channel architecture

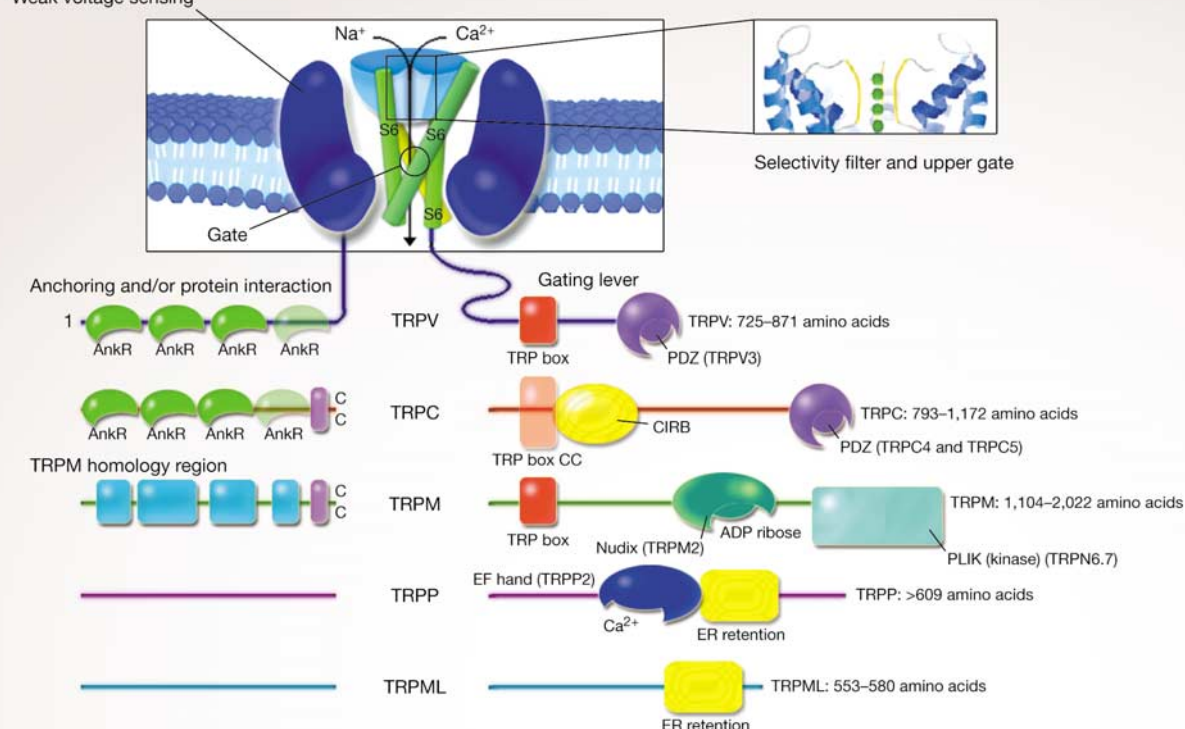
Life-forms, from bacteria to all higher plants and animals, use 6TM (S1–S6)-containing ion channels to sense and respond to stimuli. The 6TM element is one of four similar or identical subunits presumed, by analogy to K^{+} channels⁹⁷, to surround the central pore. The gate and selectivity filter are formed by the four 2TM elements (S5–pore loop–S6) facing the centre of the channel. Cations are selected for permeation by the extracellular-facing pore loop, held in place by the S5 and S6 α -helices. All TRP channels are nonselective with $P_{\text{Ca}}/P_{\text{Na}} \leq 10$, with the exception of the monovalent-selective TRPM4 and TRPM5, and the Ca^{2+} -selective TRPV5 and TRPV6.

The cytoplasmic ends of the S6 helices form the lower gate, which opens and closes to regulate cation entry into the channel. The selectivity filter itself may also gate. The S1–S4 domain may flex relative to S5–S6 in response to stimuli, but the paucity of positively charged arginines in TRP S4 helices indicates weak voltage sensitivity. All elements outside the S5–S6 region provide means of either subunit

association or act as linkers to elements that control gating.

Box 1 figure emphasizes the diversity of TRP cytoplasmic domains. The selectivity filter (light blue and inset) is formed by amino acids that dip into the bilayer (pore loops), one contributed from each of the four subunits. S5 has been removed to emphasize the link between the S6 gating helix and the TRP C-terminal polypeptide chain. The TRP box is EWKFAR in TRPC, but is less conserved in TRPV and TRPM. CC indicates a coiled-coil domain. Ankyrin repeats (AnkR) range from 0 to 14 in number (3 or 4 in TRPV and TRPC, 14 in ANKTM; not shown). Numbers on the right indicate range in length. CIRB, putative calmodulin- and IP_3R -binding domain; EF hand, canonical helix–loop–helix Ca^{2+} -binding domain; PDZ, amino acids binding PDZ domains; PLIK, phospholipase-C-interacting kinase, an atypical protein kinase intrinsic to the TRPM6 and TRPM7 polypeptide chains; Nudix, NUDT9 hydrolase protein homologue binding ADP ribose. The function of the TRPM homology region is not known. Domains not drawn to scale.

S1–S4 transmembrane domains
Weak voltage sensing



orthologue encodes an important sensor localized to neuronal microvilli in the vomeronasal organ²³. Trpc2-deficient mice display abnormal mating behaviour, consistent with a role for this channel in pheromone signalling⁷.

The TRPV (vanilloid receptor, osm9-like) subfamily

TRPV1 was identified by expression cloning using the 'hot' pepper-derived vanilloid compound capsaicin as a ligand. TRPV1 is a Ca^{2+} -permeant channel^{24,25} that is potentiated by heat ($>43^\circ\text{C}$) and decreased pH, and inhibited by intracellular phosphatidylinositol-4,5-bisphosphate (PIP_2)^{24,26}. Its thermal sensitivity is enhanced by bradykinin and nerve growth factor, which appear to act via phospholipase C (PLC) to hydrolyse PIP_2 , releasing inhibition of the channel²⁶. $\text{Trpv1}^{-/-}$ mice are defective in nociceptive, inflammatory and hypothermic responses to vanilloid compounds, supporting the interpretation that TRPV1 contributes to acute thermal nociception and hyperalgesia after tissue injury²⁷. TRPV1 also participates in mechanically evoked purinergic signalling by the bladder urothelium²⁸.

TRPV2, which is 50% identical to TRPV1 (ref. 29), may mediate high-threshold ($>52^\circ\text{C}$) noxious heat sensation, perhaps through lightly myelinated A δ nociceptors. Interestingly, TRPV2 translocates from intracellular pools upon insulin growth factor stimulation of transfected cells³⁰. Stretch reportedly increases TRPV2

translocation, and cardiac-specific transgene expression of TRPV2 results in Ca^{2+} -overload-induced cardiomyopathy³¹. But it is not surprising that overexpression of a Ca^{2+} -permeant channel induces cardiomyopathy, because such TRP channels are deleterious to many cells, including neurons. In fact, the mechanism of pain-relieving topical capsaicin is due, in part, to neuronal cell death.

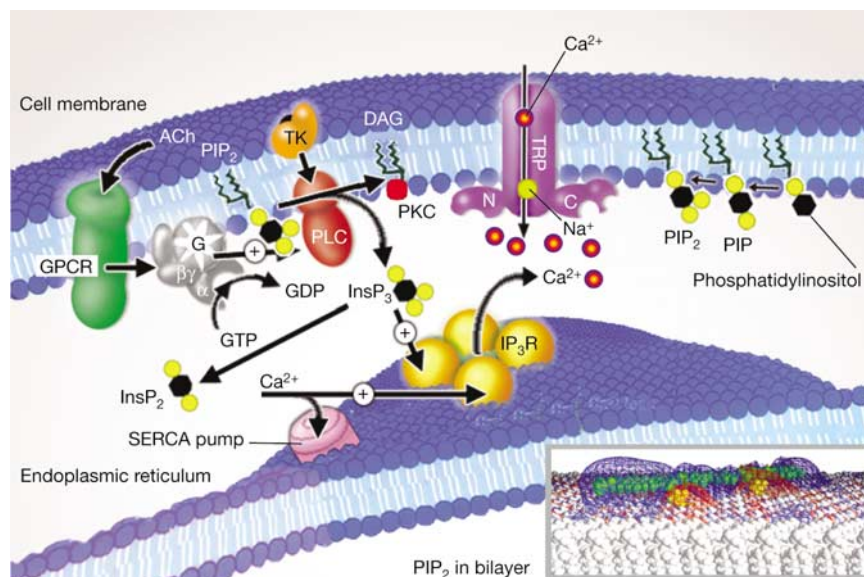
Increased temperature also activates TRPV3 ($>31^\circ\text{C}$)^{32–34} and TRPV4 ($>25^\circ\text{C}$)³⁵. The neuronal distribution of TRPV3 overlaps with TRPV1, raising the interesting possibility that they may heteromultimerize³². TRPV3 is also highly expressed in skin, tongue and the nervous system, possibly explaining the activity of 'warm-sensitive' neurons. The effect of temperature on rates of biological processes is expressed as the 10°C temperature coefficient¹: $Q_{10} = \text{rate}(T + 10^\circ\text{C})/\text{rate}(T)$. Most ion channels and enzymes have gating Q_{10} values of 3–5, but the Q_{10} of TRPV3 gating is >20 (ref. 33), and for TRPV1 and TRPV4 it is estimated to be ~ 10 – 20 (ref. 36). TRPV4 current is potentiated by hypotonicity (cell swelling)^{37,38}. $\text{Trpv4}^{-/-}$ mice have a marginally impaired renal response to hypertonicity, probably due to abnormal central control of antidiuretic hormone secretion³⁹. Hypotonicity increases TRPV4-mediated current in primary afferent nociceptive nerve fibres, an effect that is enhanced by the hyperalgesic inflammatory mediator prostaglandin E_2 (ref. 40). Expressed TRPV4 may be gated by epoxyeicosatrienoic acids⁴¹.

Box 2

Canonical TRP signal transduction

TRP channels are activated primarily by signal transduction pathways. A common pathway that is well established for *Drosophila* TRP activation is outlined in the Box 2 figure (for a review, see ref. 11). In mammalian cells, a GPCR (for example, muscarinic M1 acetylcholine receptor) catalyses G protein nucleotide exchange to form active $\text{G}\alpha$ and $\text{G}\beta\gamma$ subunits, in turn activating $\text{PLC}\beta$. Alternatively, tyrosine kinase (TK) receptors activate $\text{PLC}\gamma$. PLC hydrolyses an abundant membrane component, PIP_2 , into soluble and lipophilic messengers. Diacylglycerol, one product of PIP_2 hydrolysis, remains in the membrane. Soluble InsP_3 activates the IP_3R on the endoplasmic reticulum to release intracellular Ca^{2+} . The inset to the Box 2 figure shows PIP_2 (yellow) with its charged head group protruding above the bilayer (cytoplasmic side shown). A hydrophobic peptide (green) with interspersed basic amino acid residues sequesters PIP_2 . The $+25\text{ mV}$ electrostatic potential for the peptide (blue lines) and the -25 mV electrostatic potential for PIP_2 (red lines) are shown⁸² (inset courtesy of S. McLaughlin, SUNY Stony Brook, USA).

For a cell at rest, $[\text{Ca}^{2+}]$ is $\sim 20,000$ times lower in the cytoplasm than outside the cell. It is maintained at 50 – 100 nM concentrations by transporters and a wealth of binding proteins. Readily bound by proteins, Ca^{2+} is inherently a localized second messenger. Ca^{2+} escapes the grasp of negatively charged domains for only microseconds before being rebound, dramatically decreasing its effective diffusion coefficient. It is likely that it modulates, either directly or indirectly, all TRP channels. DAG, a diester of glycerol and two fatty acids, is best known for its anchoring and activation of protein kinase C, but it also binds and translocates other proteins (for example, RasGRPs, Munc13s and DAG kinase γ). DAG kinase (DAGK) phosphorylates DAG to form phosphatidic acid. Several TRP channels are activated by DAG and *Drosophila* TRP channels are constitutively active in DAGK-defective mutants³⁸. DAG is also converted into arachidonic acid by DAG lipase. Arachidonic acid, itself a second messenger, is the wellspring of a large cascade of active molecules.



TRPV5 and TRPV6 comprise a distinct subfamily of homomeric and heteromeric channels found in transporting epithelia of the kidney and intestine. They show strong inwardly rectifying currents and are the most Ca^{2+} -selective TRP channels (permeability ratio $P_{\text{Ca}}/P_{\text{Na}} > 100$)^{42,43}, suggesting that they mediate Ca^{2+} uptake. Both are inactivated by $[\text{Ca}^{2+}]_i$ (ref. 44); TRPV6 shows voltage-dependent intracellular Mg^{2+} blockade⁴⁵.

The TRPM (melastatin) subfamily and TRPA

TRPM1 (melastatin) was initially identified as a transcript that showed decreased expression in highly metastatic versus non-metastatic melanoma cells⁴⁶. It is widely expressed in normal tissues, but its function and electrophysiological properties have not been described.

TRPM2 (ref. 47) forms a Ca^{2+} -permeant channel⁴⁸ that is gated by binding of ADP ribose ($\text{EC}_{50} \approx 100 \mu\text{M}$) and nicotinamide adenine dinucleotide (NAD; $\sim 1 \text{ mM}$)^{48,49} to a carboxy-terminal NUDT9 Nudix hydrolase domain. ADP ribose is a breakdown product of NAD, CD38, cyclic ADP ribose (a Ca^{2+} -release messenger) and protein de-acetylation (*O*-acetylated ADP ribose⁵⁰), but the TRPM2 domain itself is an ineffective hydrolase⁵¹. The channel is regulated by signalling pathways responsive to H_2O_2 and tumour-necrosis factor- α , suggesting that it may act as a sensor of intracellular oxidation/reduction⁵², possibly during the oxidative burst of neutrophils⁵³.

TRPM3 (refs 54, 55) forms a Ca^{2+} -permeant nonselective channel that is constitutively active when heterologously expressed. Its activity is increased by hypotonicity (200 mOsm per litre), but there is little homology to TRPV4 that might suggest a common mechanism of activation⁵⁴. TRPM4 is expressed primarily in kidney

distal-collecting-duct epithelium and in the central nervous system.

TRPM4 and TRPM5 are the only monovalent-selective ion channels of the TRP family. They are widely distributed and may account for observed Ca^{2+} -activated ~ 20 – 30 pS nonselective channel activities^{56,57}. They are activated through GPCRs coupled to PLC-dependent endoplasmic reticular Ca^{2+} release, perhaps by direct Ca^{2+} binding to the channel. However, relatively high $[\text{Ca}^{2+}]_i$ is required to activate these channels^{56,57}, suggesting that they localize close to sites of Ca^{2+} release or that other modulators are important. Although their instantaneous *I*–*V* relationships are linear, the fraction of open channels increases at positive potentials. This voltage dependence is not mediated by divalent cation binding, suggesting an intrinsic voltage-sensing mechanism^{57,58}.

TRPM5 is found in cells expressing taste receptors⁵⁹. In an *in vivo* study in $\text{TrpM5}^{-/-}$ mice, it was shown that taste receptors T1R and T2R share a common signalling pathway involving PLC β 2 and TRPM5, to produce sweet, umami and bitter taste sensations⁸. The authors concluded that InsP_3 , Ca^{2+} and thapsigargin-mediated store depletion did not activate TRPM5. However, it is possible that PIP_2 or other molecules modulate its sensitivity to $[\text{Ca}^{2+}]_i$.

TRPM6 and TRPM7 are unique among ion channels because they also contain functional kinase domains⁶⁰. TRPM7 passes little inward current under physiological conditions, is permeant to both Ca^{2+} and Mg^{2+} , and is inhibited by $\sim 0.6 \text{ mM}$ intracellular free Mg^{2+} (refs 61, 62). In contrast to other GPCR-activated TRP channels, TRPM7 current increases slowly under whole-cell recording conditions and is inactivated by PIP_2 hydrolysis by PLC β or PLC γ ⁶². The function of the kinase domain is poorly understood and its substrates have not been identified. In contrast to our original

Table 1 Identifying properties of TRP channel subunits

TRP	Functions	Interacting proteins	Activation	$P_{\text{Ca}}/P_{\text{Na}}$
C1 250687	Central nervous system (CNS)	TRPC4; TRPC5; TRPC3 (embryo); calmodulin; IP ₃ R; caveolin-1; G $\alpha_{q/11}$; PMCA	GPCR (G $_q$ -PLC β); TKR-PLC γ ?	<10
C2 *602343	Pheromone sensing		GPCR (G $_q$ -PLC β); TKR-PLC γ ?	
C3 150981	Vasoregulation (resistance arteries); airway regulation; CNS	TRPC1 (embryo); PLC β ; IP ₃ R; ryanodine receptor; SERCA; caveolin-1; calmodulin	GPCR (G $_q$ -PLC β); DAG; TKR-PLC γ ; $[\text{Ca}^{2+}]_i$?	1.6
C4 262960	Vasoregulation; microvascular permeability?; CNS; gastrointestinal motility?	TRPC1; TRPC5; PLC β ; NHERF via PDZ-binding domain; IP ₃ R; calmodulin	GPCR (G $_q$ -PLC β); TKR-PLC γ ?; GTP γ S; La^{3+} ; calmidazolium	~ 1
C5 247868	CNS; growth cone morphology	TRPC1; TRPC4; calmodulin; PLC β ; stathmin	GPCR (G $_q$ -PLC β); GTP γ S; La^{3+}	~ 1
C6 159003	Vasoregulation (resistance arteries); airway regulation; smooth muscle?	TRPC3; TRPC7; calmodulin	GPCR (G $_q$ -PLC β); DAG; TKR-PLC γ ?; $[\text{Ca}^{2+}]_i$; AlF_4 ; flufenamate	5
C7 283104		TRPC3; TRPC6; calmodulin	GPCR (G $_q$ -PLC β); DAG; TKR-PLC γ ?; $[\text{Ca}^{2+}]_i$	~ 0.5
V1 (VR1) 283010	'Hot' pepper and heat sensation; inflammatory thermal hyperalgesia; bladder distension sensation	TRPV3; calmodulin	$T > 43^\circ\text{C}$; capsaicin; resiniferatoxin; anandamide; H^+	10 ($P_{\text{Mg}}/P_{\text{Na}} \approx 5$)
V2 (VRL1; GRC) 279746	Noxious heat sensing; muscle degeneration?; other pain pathways		$T > 53^\circ\text{C}$; insulin growth factor-1	3
V3 375034	Warm sensing?; pain?	TRPV1	$T > 30^\circ\text{C}$	~ 10
V4 (OTRPC4; VR-OAC) 287776	CNS osmotic sensing; pressure sensing (dorsal root ganglion); nociception; warm sensing	Calmodulin; Src family kinases	Hypo-osmolarity; $T > 24^\circ\text{C}$; phorbol esters; anandamide; arachidonic acid; epoxyeicosatrienoic acid	6
V5 (ECaC1; CaT2) 283369	Ca^{2+} uptake in kidney; intestine?	TRPV6; S100A10/annexin 2	Constitutive	>100
V6 (ECaC2; CaT1) 302740	Ca^{2+} uptake in intestine?	TRPV5; S100A10/annexin 2	Constitutive	>100
M1 (MLSN) 43265				
M2 133517	Oxidant stress sensor?		ADP ribose; βNAD ; H_2O_2 ?	~ 1 (estimate)
M3 288911	Ca^{2+} uptake in kidney?		Constitutive; hypo-osmolarity $[\text{Ca}^{2+}]_i$	1.6
M4 31608			$[\text{Ca}^{2+}]_i$	<0.05
M5 (Mtr1) 272287	Taste (sweet; bitter; umami)		T1R; T2R-G $_{\text{gus}}$ -PLC β 2; $[\text{Ca}^{2+}]_i$	<0.05
M6 (CHAK2) 272225	Mg^{2+} uptake in kidney and intestine?			
M7 (TRP-PLIK) 33819	Cellular Mg^{2+} homeostasis?	PLC β 1; PLC β 2; PLC β 3; PLC γ 1	PIP_2 ; tyrosine phosphate	0.3
M8 (CMR1) 366053	Noxious cold sensing; nociception; cancer?		$T < 25^\circ\text{C}$; menthol and icilin potentiate	1–3
A1 (ANKTM1) 186329	Noxious cold sensing		$T < 18^\circ\text{C}$; icilin potentiates	1.4 ($P_{\text{Ca}} \approx P_{\text{Mg}}$)
P2 (PC2; PKD2) 82001	Mechanosensing in cilia?; fertility?	PC1; Hax-1; cortactin	Mechanical stress?; $[\text{Ca}^{2+}]_i$?	~ 1 –5
P3 (PKD2L1) 159241	Kidney and retinal development?		$[\text{Ca}^{2+}]_i$	~ 4
P5 (PKD2L2) 272418			$[\text{Ca}^{2+}]_i$?	~ 1 –5 ($P_{\text{Mg}} \approx 0$)

Numbers in TRP column are Unigene designation (human, except for *mouse). GPCR (G $_q$ -PLC β) indicates the canonical TRP signalling cascade. For complete table and tissue distribution, see <http://clapham.tch.harvard.edu/>.

report, it is not required for channel activation^{62,63}. The catalytic core of the kinase domain is similar to that of other eukaryotic protein kinases and to enzymes with ATP-grasp domains⁶⁴. The sensitivity of TRPM7 to physiological Mg-ATP levels has been suggested to have a central role in metabolic sensing⁶¹ or to serve as a mechanism to adjust cellular Mg²⁺ homeostasis⁶³. But a spontaneous human mutation in TRPM6 results in familial hypomagnesaemia with secondary hypocalcaemia, suggesting that TRPM6 may be important for Mg²⁺ uptake in the kidney and intestine.

TRPM8 was identified as a messenger RNA that was upregulated in prostatic and other cancers. Its sensory role was recognized when it was isolated by expression cloning of a menthol receptor from trigeminal neurons^{65,66}. TRPM8 is a nonselective, outwardly rectifying channel that can be activated by cold (8–28 °C) and enhanced by ‘cooling’ compounds such as menthol and icilin. TRPM8 is widely expressed, but thought to function specifically as a thermosensor in TrkA⁺, small-diameter primary sensory neurons.

ANKTM1, a Ca²⁺-permeant, nonselective channel homologous to *Drosophila painless*⁶⁷, is distinguished by ~14 amino-terminal ankyrin repeats. It is activated by noxious cold temperature (<15 °C) but bears little similarity to menthol-sensitive TRPM8 (ref. 68). It is found in a subset of nociceptive sensory dorsal root ganglion neurons, in the company of capsaicin-sensitive TRPV1, but not TRPM8. Interestingly, the *Drosophila* orthologue of ANKTM1 responds to warming (>27 °C) rather than to cooling when expressed in *Xenopus* oocytes⁶⁹. These observations are consistent with sensitivity to the surrounding membrane environment, but might also be reconciled if the lowest energy state of the mammalian channel is the open configuration. Other TRP channels have not been systematically tested for temperature sensitivity, but such a comparison would clarify this issue.

The TRPP (polycystin) and TRPL (mucolipin) subfamilies

Polycystic kidney disease proteins PKD2, PKD2L1 and PKD2L2 are 6TM Ca²⁺-permeant channels called TRPP2, TRPP3 and TRPP5, respectively. The much larger TRPP1, polycystin-REJ and polycystin-1L1 proteins are 11TM proteins that contain a C-terminal 6TM TRP-like channel domain. TRPP1 is not known to form a channel by itself, but it complexes with TRPP2 to form a Ca²⁺-permeable nonselective cation channel⁷⁰. Autosomal dominant polycystic kidney disease is caused by mutations in TRPP1 or TRPP2, leading to alterations in the polarization and function of cyst-lining epithelial cells. *Trpp1*^{-/-} and *Trpp2*^{-/-} mice die *in utero* with cardiac septal defects and cystic changes in nephrons and pancreatic

ducts^{71,72}. The mouse orthologue of TRPP3 is deleted in *krd* mice, resulting in defects in the kidney and retina⁷³.

TRPP proteins have another role in development. Normal body asymmetry appears to arise from leftward extracellular flow generated by motor-protein-dependent rotation of monocilia on the ventral surface of the embryonic node. Motile monocilia generate nodal flow, and non-motile TRPP2-containing cilia sense nodal flow, initiating an asymmetric Ca²⁺ signal at the left nodal border⁷⁴. TRPP1 and TRPP2 both appear to be targeted to primary cilia cells of renal epithelia, where the channel complex is gated by fluid flow⁷⁵.

The mucolipins (MCOLN1, MCOLN2 and MCOLN3) are 6TM channels that are probably restricted to intracellular vesicles. Mutations in MCOLN1 (TRPL1) are associated with mucopolidosis type IV, a neurodegenerative lysosomal storage disorder⁷⁶. The defect appears to be in sorting or transport in the late endocytic pathway. Mutations in a *Caenorhabditis elegans* TRPL1 homologue, *cup-5*, cause excess lysosome formation and apoptosis in all cell types⁷⁷. TRPL3 is present in the cytoplasm of hair cells and the plasma membrane of stereocilia. TRPL3 is mutated in the *varitint-waddler* mouse, resulting in deafness and pigmentation defects⁷⁸.

Theories of TRP channel gating

The gnawing mystery of TRP channels is their elusive mechanism of gating. Many TRP channels show some constitutive activity in overexpression systems, but only a few have been studied in their native environment or at physiological temperatures. Our poor understanding of TRP channel gating may reflect our ignorance of potential intra- or extracellular ligands. But a more interesting possibility is that there might be a common TRP gating mechanism. Attempts to understand TRP channel activation have given rise to several theories, discussed below.

Receptor-operated theory

This is the most likely mechanism for TRPC channels and the *Drosophila* photoreceptor TRP (Box 2). All mammalian TRPC channels can be activated by GPCRs. These include muscarinic type 1 (TRPC1, TRPC4, TRPC5 heteromers, or presumed TRPC4, TRPC5 homomers), histaminergic type 1 (TRPC3, TRPC6) and purinergic receptors (TRPC7). GPCRs are often attached to multi-molecular complexes by scaffolding proteins, adding an extra level of complexity. To date, investigators have identified only the proteins that interact prior to PLC activation. Unfortunately, PLC activation generates DAG and free cytoplasmic Ca²⁺ (via InsP₃), both of which can unleash legions of active signal transduction molecules. The major challenge of the receptor-operated theory is to find a receptor-activated messenger that directly binds and specifically activates a TRP channel. To establish this theory for a subset of TRP channels, it will be necessary to identify the native receptors that activate a particular channel, the downstream messengers mediating activation, and the physiological roles of these proteins and messengers in the context of the tissues that normally express them. One potential messenger, PIP₂, is described in more detail.

PIP₂ comprises ~1% of anionic phospholipids in cells and is much more abundant than its signalling relative PIP₃. PIP₂ regulates at least seven types of ion channel and transporter⁷⁹. Because several TRPs (TRPC3, TRPC4, TRPM7) are known to bind PLCβ and/or PLCγ, PLC in part determines the PIP₂ concentration that TRPs encounter. PIP₂ inhibits *Drosophila* TRP and TRPL⁸⁰ and mammalian TRPV1 (ref. 26), but increasing PIP₂ lowers the temperature activation threshold of TRPV1 (ref. 81) and probably other TRP channels. Constitutive TRPM7 activity is increased by PIP₂ and inactivated by PIP₂ hydrolysis at 22 °C (ref. 62). These results suggest that PIP₂ commonly interacts with positively charged regions of ion channels to alter the energy required for gating. Is PIP₂ a common denominator of TRP activation?

The head group of PIP₂, carrying about four negative charges, may

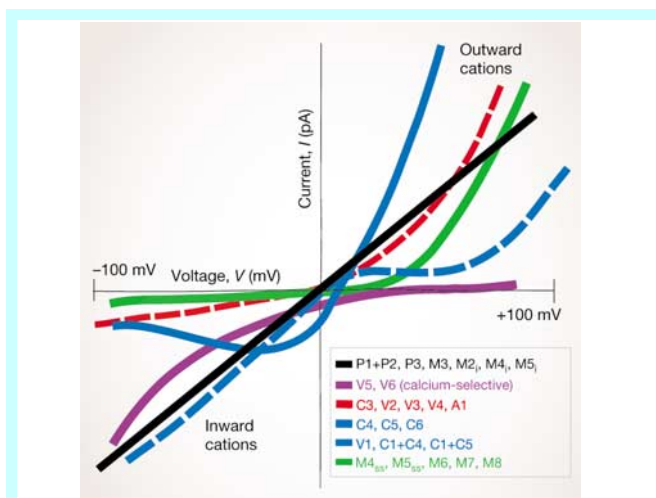


Figure 2 Representative transmembrane currents flowing in response to a set voltage ($-V$ relation) of various TRP channels. Inward cation flow is described as negative current. ss, steady state; i, instantaneous.

'snorkel' into the aqueous phase above other phospholipids (inset to Box 2 figure). Positively charged peptides containing basic residues can hug the bilayer if they contain aromatic amino acids (particularly Phe and Trp). If the peptide also contains clusters of basic residues (≥ 4), it can electrostatically fence in PIP_2 , forming basins of concentrated PIP_2 (ref. 82). Indeed, such a cluster in the mid-C terminus is required for inhibition of TRPV1 by PIP_2 (ref. 81), but is missing in the PIP_2 -insensitive TRPV3 (S. Ramsey, unpublished work). Interestingly, all TRPs contain a Trp/Phe segment with basic lysine and arginine residues (TRP box) just distal to the S6 gating helix/cytoplasmic junction, but there may not be enough positive charge in this region to sequester PIP_2 . Considering the data to date, PIP_2 is likely to modulate gating of some TRP channels, but it is not a unifying mechanism of TRP channel activation.

Store-operated calcium entry hypothesis

Putney proposed that emptied Ca^{2+} stores (primarily in the endoplasmic reticulum) somehow gate the entry of external Ca^{2+} to replenish the deficit (for a review, see ref. 83). This mechanism is called capacitative Ca^{2+} entry, or store-operated Ca^{2+} entry (SOCE). The physiological hallmark of SOCE is a large, receptor-mediated, transient $[\text{Ca}^{2+}]_i$ increase followed by a prolonged high $[\text{Ca}^{2+}]_i$ plateau that is dependent on $[\text{Ca}^{2+}]_o$ (ref. 84). Thapsigargin, an inhibitor of smooth endoplasmic reticulum Ca^{2+} -ATPase (SERCA) pumps, is often used to examine this phenomenon, with the assumption that the drug is specific for SERCAs. The SOCE hypothesis has intrigued many scientists and has led to the search for an ion channel and a mechanism that could link store depletion with Ca^{2+} entry. I_{CRAC} (calcium-release-activated current) is the best candidate for a store-depletion-responsive current, but it does not appear to be mediated by any TRP protein. None of the TRP channels exhibits the requisite high Ca^{2+} selectivity ($P_{\text{Ca}}/P_{\text{Na}} \approx 1,000$), low single-channel conductance (< 0.1 pS), and pharmacological enhancement by 1–5 μM 2-APB⁸⁵. The three main theories for SOCE are a direct coupling mechanism, a diffusible messenger, and store-depletion-mediated fusion of a vesicle containing a Ca^{2+} -permeant channel (Fig. 3).

Although dozens of papers tout the link between SOCE and TRP channels, it is worth re-examining the data. The most common

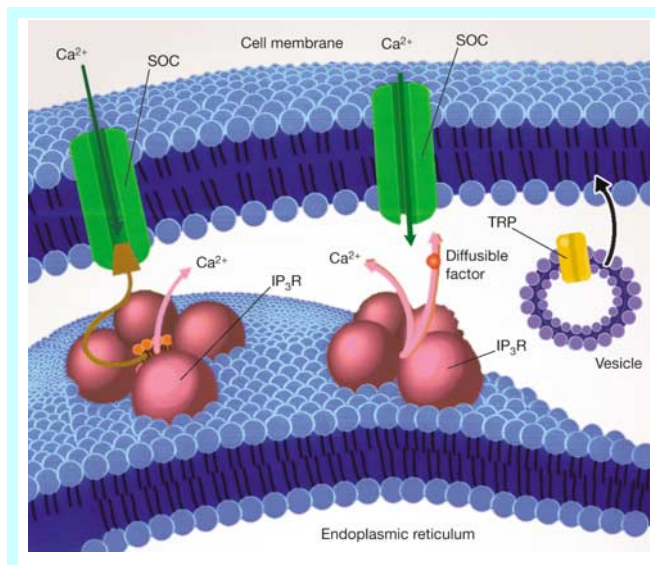


Figure 3 Three theories of store-operated Ca^{2+} entry. Depletion of Ca^{2+} in the endoplasmic reticulum causes opening of the SOC. Left, a direct link between the IP_3R opens such a channel, in analogy to the link between ryanodine receptors and Ca_v channels. Middle, a diffusible second messenger released from Ca^{2+} -depleted endoplasmic reticulum activates the SOC. Right, Ca^{2+} depletion results in fusion of vesicles containing SOCs.

standard of proof has been to show, either by Ca^{2+} imaging or by electrophysiology, that a particular heterologously expressed TRP channel enhances store-depletion-mediated entry. However, a better standard for this activity would be total abrogation of the phenomenon by elimination of a protein or gene, either *in vivo* or in a faithful model of the native system.

Nilius and colleagues¹⁸ studied $\text{Trpc4}^{-/-}$ mice and concluded that TRPC4 was an essential component of SOCE channel activation. However, electrophysiological characterization suggested that the defect resulted from a decrease in I_{CRAC} , and TRPC4 alone has few properties characteristic of CRAC channels. Our laboratory suggested that the Ca^{2+} -selective TRPV6 might be a component of a CRAC channel⁴², and a dominant-negative TRPV6 subunit in Jurkat cells was interpreted as suppressing endogenous I_{CRAC} ⁸⁶. But TRPV6 alone does not account for several properties of I_{CRAC} ^{42,87}. In a careful Ca^{2+} imaging and electrophysiological study using double-stranded RNA to target endogenous TRPC1-containing channels in Chinese hamster ovary cells, it was concluded that TRPC1 is an essential component of SOCE⁸⁸. Despite the use of knockout or knockdown methods, none of these proteins is universally accepted as a store-operated Ca^{2+} channel. In my opinion, the major molecules comprising CRAC remain unknown.

Almost every publication concluding that a TRP is a SOC is countered by a paper stating that the same TRP is not a SOC. Most of the data based on Ca^{2+} imaging are pro-SOCE, whereas most of the electrophysiological data are con-SOCE. Ca^{2+} imaging is prone to false-positives and electrophysiology is prone to false-negatives. Ca^{2+} imaging has an advantage in that cells can be kept in a fairly normal environment. But Ca^{2+} imaging is a very indirect assay of channel function because it reflects accumulated free $[\text{Ca}^{2+}]_i$, regardless of the source. Addition of thapsigargin without accounting for constitutive Ca^{2+} entry leads to falsely high $[\text{Ca}^{2+}]_i$. SERCA pump inhibition can also artificially increase the Ca^{2+} signal by eliminating the endoplasmic reticulum as a Ca^{2+} buffer and reducing the effective Ca^{2+} volume of the cell. Ca^{2+} entry into the unbuffered cytoplasm may induce Ca^{2+} -activated channels, Ca^{2+} transporters, or other unrelated Ca^{2+} -dependent processes that affect the results (this problem can be reduced by substitution of extracellular Ca^{2+} with Ba^{2+}). Most importantly, voltage levels, the driving force for Ca^{2+} entry into the cytoplasm, are uncontrolled.

Patch clamp is a direct assay of channel activity that overcomes most of these limitations. In patch clamp, the channel can be bathed in defined solutions both inside and outside the cell, and voltage is controlled. Excellent time and current resolution allows the dissection of independent processes. Test compounds and proteins can readily be applied to either side of the channel. However, in standard patch-clamp recordings, intracellular contents are perfused, rapidly replacing small molecules and diffusible proteins. This problem can be circumvented by perforated patch recording, in which molecules larger than ions are restricted. Unfortunately, although it is not difficult to perform, it has rarely been used for these studies. Finally, electrophysiologists routinely record at 22 °C, and the few TRP channels that have been tested are highly sensitive to temperature. These problems could easily be rectified, and if they are, electrophysiology will most probably become the standard for SOC identification.

In my opinion, Ca^{2+} -permeant TRPs, like all Ca^{2+} -permeant channels, contribute to $[\text{Ca}^{2+}]_i$ and thus affect the SOCE process. But they have not been proved to be SOCs in any direct sense. The SOCE hypothesis for TRPs will not be settled until there is a consensus on the molecular identity of SOCs and the endoplasmic-reticulum-dependent signal that activates them. In the interim, the use of the term 'store-operated channel' for TRPs is confusing and should be avoided.

Vesicle fusion hypothesis

A Ca^{2+} entry channel was proposed to fuse with the plasma membrane to mediate SOCE, but neither the channel mediating

this event nor the mechanism was identified^{89,90}. Independent of the SOCE theory, do TRPs rapidly translocate to the plasma membrane in response to stimuli? The mucolipins (TRPML) are involved in intravesicular trafficking, but little more is known of their function or whether they can also be present in the plasma membrane of native cells. Interestingly, TRPC1 and TRPC3–TRPC6 are all present in rat brain synaptosomes⁹¹. Recently, *C. elegans* sperm TRP (TRP-3), required for sperm–egg interaction, was localized in intracellular vesicles until fertilization competence required TRP-3 translocation from the vesicles to the plasma membrane⁹².

These observations raise the question of whether some TRP channels are held in reserve and are then rapidly placed in the cell membrane in response or in adaptation to a stimulus. If so, what mechanism links the stimulus to the translocation of TRP-containing vesicles? And are the TRP channels in vesicles required for vesicle trafficking, swelling and/or the fusion process itself?

Cell sensory hypothesis

Block⁹³ beautifully encapsulated the underlying physical principles of sensory transduction. At the heart of sensation is the ability to distinguish input from a photon (vision, heat, electromagnetic force), phonon (sound), chemical (for example, odorant), or mechanical force (stretch, osmolarity, gravity) from background thermal energy or extraneous inputs (noise). Evolution has enhanced these abilities not just by sensitive detection, but also by amplification and signal processing. Built into almost all TRPs is the ability to conduct Ca^{2+} ions in a parsimonious manner; their I – V relations, nonselective character and low density dictate that they admit relatively little Ca^{2+} per second compared with, for example, a voltage-gated Ca^{2+} channel (Ca_V). Second, they are active at resting membrane potentials where cells spend most of their time. These characteristics can be harmful to a cell if TRPs are overexpressed, as is commonly observed in expression systems or in some cancer cells. As for most Ca^{2+} -permeant channels, the majority of TRPs are inherently self-inactivating by virtue of their $[\text{Ca}^{2+}]_i$ sensitivity. These properties provide a first level of signal processing common to sensory mechanisms, but what about detection?

Sensation requires the detection of force. The primary mechanisms for TRP activation involve mechanical force, intracellular ligand binding (signal transduction molecules) and temperature. How do these forces translate into channel gating? The Boltzmann equation gives the equilibrium distribution of channels in the open and closed states; the fraction of open channels is $[1 + \exp(\Delta G/Nk_B T)]^{-1}$, where ΔG is the free energy of transition between the closed and open state, k_B is Boltzmann's constant, N is Avogadro's number and T is the absolute temperature. The free energy, ΔG , is equal to $w - \psi$, where w is the conformational energy increase upon gating the channel and ψ represents the sum of energies transduced to the channel from external forces. Voltage-sensitive channels are usually held in a higher-energy, closed state at resting membrane potentials; removal of this energy (depolarization to 0 mV) allows the channel protein to relax into its lower-energy structural configuration. To sense this electromotive energy, voltage-sensitive ion channels have charged amino acid 'solenoids' that drive conformational changes. Most TRP channels lack these attributes and at the resting membrane potential are contentedly immune to the transmembrane field. What then is the ψ for TRPs?

The detection limit for an open TRP channel is imposed by thermal energy ($\sim 0.6 \text{ kcal mol}^{-1}$ at 37°C). This energy is tiny, less than the energy of a single photon of visible light per molecule. But other cellular noise sources must be overcome, probably requiring channel gating for the cell to record a significant stimulus. The gating energy for the membrane-tension-sensitive bacterial channel in a liposome is $\sim 10 \text{ kcal mol}^{-1}$ (ref. 94), only 20 times the thermal limit. Mechanosensation—the basis of hearing, osmolar sensing, stretch and flow sensing—easily provides these levels of energy, especially if the mechanical advantage provided by cilia is taken into

account⁹⁵. TRPC2 (ref. 23), OSM-9 (ref. 6), TRPP2 (PKD)⁷⁵ and *Nan*⁹⁶ are all found in ciliated structures. TRPs are also common within ciliated cells that sense flow in kidney epithelia, vascular endothelia, lung and intestine, and in the hearing and vestibular apparatus, taste cells and odorant-sensing cells. Anchoring TRPs to mechanical forces is not without its risks to the cell. Prolonged Ca^{2+} entry due to abnormal membrane tension, such as occurs with loss of membrane-stabilizing proteins in muscular dystrophy, may result in muscle degeneration. Prolonged stretch in cardiac muscle leads to Ca^{2+} overload, hypertrophy and cardiac failure.

For signal-transduction-gated TRPs, the energies imparted by intracellular ligand binding are also sufficient for gating, depending on the binding affinity of the specific interaction (dissociation constant values of $1 \mu\text{M}$ imply ~ 1 – 10 kcal mol^{-1}). But the most obscure mechanism for activation of TRPs is temperature. Oddly, all TRP recordings published to date fail to reach steady state and do not saturate within the range of practical recordings. This prevents estimates of gating energy, but temperature-sensitive TRP channels have estimated Q_{10} values up to ten times larger than is typical for enzymes or channels. Temperature-dependent TRPV3 current is reversible upon cooling, but exhibits a pronounced hysteresis; the current increases with increasing temperature but abruptly collapses when cooling begins³³. Perhaps heat induces lipid bilayer rearrangements to alter membrane tension. Alternatively, the protein may denature (melt) and rapidly refold when energy is removed. Finally, TRPs could be gated by temperature-dependent cooperative binding by second messengers.

What's next?

The large number of TRP subtypes, their overlapping electrophysiological characteristics, broad expression patterns, heteromultimerization, lack of specific blockers, and poorly understood mechanisms of activation have made their study difficult for TRP channel researchers. But the field is now progressing beyond simple overexpression and Ca^{2+} imaging, and more specific tools are being used. In the near future, a wealth of data from genetically altered animals will emerge, as well as useful molecular tools and assays. The next few years should determine whether any of the TRPs are indeed SOC, or whether this term has simply substituted for a poorly understood activation mechanism. TRPs are sensory molecules, but not just in the usual organismal definition of sensation. TRP channels are intrinsic sensors of the cellular environment. □

doi:10.1038/nature02196.

- Hille, B. *Ion Channels of Excitable Membranes* 3rd edn, 663–723 (Sinauer, Sunderland, MA, 2001).
- Minke, B. *Drosophila* mutant with a transducer defect. *Biophys. Struct. Mech.* **3**, 59–64 (1977).
- Montell, C., Jones, K., Hafen, E. & Rubin, G. Rescue of the *Drosophila* phototransduction mutation *trp* by germline transformation. *Science* **230**, 1040–1043 (1985).
- Denis, V. & Cyert, M. S. Internal Ca^{2+} release in yeast is triggered by hypertonic shock and mediated by a TRP channel homologue. *J. Cell Biol.* **156**, 29–34 (2002).
- Zhou, X. L. et al. The transient receptor potential channel on the yeast vacuole is mechanosensitive. *Proc. Natl Acad. Sci. USA* **100**, 7105–7110 (2003).
- de Bono, M., Tobin, D. M., Davis, M. W., Avery, L. & Bargmann, C. I. Social feeding in *Caenorhabditis elegans* is induced by neurons that detect aversive stimuli. *Nature* **419**, 899–903 (2002).
- Stowers, L., Holy, T. E., Meister, M., Dulac, C. & Koentges, G. Loss of sex discrimination and male–male aggression in mice deficient for TRP2. *Science* **295**, 1493–1500 (2002).
- Zhang, Y. et al. Coding of sweet, bitter, and umami tastes: different receptor cells sharing similar signaling pathways. *Cell* **112**, 293–301 (2003).
- Harteneck, C., Plant, T. D. & Schultz, G. From worm to man: three subfamilies of TRP channels. *Trends Neurosci.* **23**, 159–166 (2000).
- Clapham, D. E., Runnels, L. W. & Strübing, C. The TRP ion channel family. *Nature Rev. Neurosci.* **2**, 387–396 (2001).
- Minke, B. & Cook, B. TRP channel proteins and signal transduction. *Physiol. Rev.* **82**, 429–472 (2002).
- Montell, C., Birnbaumer, L. & Flockerzi, V. The TRP channels, a remarkably functional family. *Cell* **108**, 595–598 (2002).
- Birnbaumer, L., Yidirim, E. & Abramowitz, J. A comparison of the genes coding for canonical TRP channels and their M, V and P relatives. *Cell Calcium* **33**, 419–432 (2003).
- Nilius, B. TRP Channels: facts, fiction, challenges. *Cell Calcium* **14**, 33–34 (2003).
- Corey, D. New TRP channels in hearing and mechanosensation. *Neuron* **39**, 585–588 (2003).
- Wes, P. D. et al. TRPC1, a human homolog of a *Drosophila* store-operated channel. *Proc. Natl Acad. Sci. USA* **92**, 9652–9656 (1995).
- Greka, A., Navarro, B., Oancea, E., Duggan, A. & Clapham, D. E. TRPC5 is a regulator of hippocampal

- neurite length and growth cone morphology. *Nature Neurosci.* **6**, 837–845 (2003).
- Freichel, M. *et al.* Lack of an endothelial store-operated Ca^{2+} current impairs agonist-dependent vasorelaxation in $\text{TRP4}^{-/-}$ mice. *Nature Cell Biol.* **3**, 121–127 (2001).
- Tiruppathi, C. *et al.* Impairment of store-operated Ca^{2+} entry in $\text{TRPC4}^{-/-}$ mice interferes with increase in lung microvascular permeability. *Circ. Res.* **91**, 70–76 (2002).
- Zitt, C. *et al.* Expression of TRPC3 in Chinese hamster ovary cells results in calcium-activated cation currents not related to store depletion. *J. Cell Biol.* **138**, 1333–1341 (1997).
- Trebak, M., Vazquez, G., Bird, G. S. & Putney, J. W. The TRPC3/6/7 subfamily of cation channels. *Cell Calcium* **33**, 451–461 (2003).
- Inoue, R. *et al.* The transient receptor potential protein homologue TRP6 is the essential component of vascular α_1 -adrenoceptor-activated Ca^{2+} -permeable cation channel. *Circ. Res.* **88**, 325–332 (2001).
- Liman, E. R., Corey, D. P. & Dulac, C. TRP2: a candidate transduction channel for mammalian pheromone sensory signaling. *Proc. Natl Acad. Sci. USA* **96**, 5791–5796 (1999).
- Caterina, M. J. *et al.* The capsaicin receptor: a heat-activated ion channel in the pain pathway. *Nature* **389**, 816–824 (1997).
- Jordt, S. E. & Julius, D. Molecular basis for species-specific sensitivity to “hot” chili peppers. *Cell* **108**, 421–430 (2002).
- Chuang, H. H. *et al.* Bradykinin and nerve growth factor release the capsaicin receptor from PtdIns(4,5) P_2 -mediated inhibition. *Nature* **411**, 957–962 (2001).
- Caterina, M. J. *et al.* Impaired nociception and pain sensation in mice lacking the capsaicin receptor. *Science* **288**, 306–313 (2000).
- Birder, L. A. *et al.* Altered urinary bladder function in mice lacking the vanilloid receptor TRPV1. *Nature Neurosci.* **5**, 856–860 (2002).
- Caterina, M. J., Rosen, T. A., Tominaga, M., Brake, A. J. & Julius, D. A capsaicin-receptor homologue with a high threshold for noxious heat. *Nature* **398**, 436–441 (1999).
- Kanzaki, M. *et al.* Translocation of a calcium-permeable cation channel induced by insulin-like growth factor-I. *Nature Cell Biol.* **1**, 165–170 (1999).
- Iwata, Y. *et al.* A novel mechanism of myocyte degeneration involving the Ca^{2+} -permeable growth factor-regulated channel. *J. Cell Biol.* **161**, 957–967 (2003).
- Smith, G. D. *et al.* TRPV3 is a temperature-sensitive vanilloid receptor-like protein. *Nature* **418**, 186–190 (2002).
- Xu, H. *et al.* TRPV3 is a calcium-permeable temperature-sensitive cation channel. *Nature* **418**, 181–186 (2002).
- Peier, A. M. *et al.* A heat-sensitive TRP channel expressed in keratinocytes. *Science* **296**, 2046–2049 (2002).
- Guler, A. D. *et al.* Heat-evoked activation of the ion channel, TRPV4. *J. Neurosci.* **22**, 6408–6414 (2002).
- Benham, C. D., Gunthorpe, M. J. & Davis, J. B. TRPV channels as temperature sensors. *Cell Calcium* **33**, 479–487 (2003).
- Strotmann, R., Harteneck, C., Nunnenmacher, K., Schultz, G. & Plant, T. D. OTRPC4, a nonselective cation channel that confers sensitivity to extracellular osmolarity. *Nature Cell Biol.* **2**, 695–702 (2000).
- Liedtke, W. *et al.* Vanilloid receptor-related osmotically activated channel (VR-OAC), a candidate vertebrate osmoreceptor. *Cell* **103**, 525–535 (2000).
- Mizuno, A., Matsumoto, N., Imai, M. & Suzuki, M. Impaired osmotic sensation in mice lacking TRPV4. *Am. J. Physiol. Cell Physiol.* **285**, C96–C101 (2003).
- Alessandri-Haber, N. *et al.* Hypotonicity induces TRPV4-mediated nociception in rat. *Neuron* **39**, 497–511 (2003).
- Watanabe, H. *et al.* Anandamide and arachidonic acid use epoxyeicosatrienoic acids to activate TRPV4 channels. *Nature* **424**, 434–438 (2003).
- Yue, L., Peng, J. B., Hediger, M. A. & Clapham, D. E. CaT1 manifests the pore properties of the calcium-release-activated calcium channel. *Nature* **410**, 705–709 (2001).
- den Dekker, E., Hoenderop, J. G., Nilius, B. & Bindels, R. J. The epithelial calcium channels, TRPV5 and TRPV6: from identification towards regulation. *Cell Calcium* **33**, 497–507 (2003).
- Hoenderop, J. G. *et al.* Homo- and heterotetrameric architecture of the epithelial Ca^{2+} channels TRPV5 and TRPV6. *EMBO J.* **22**, 776–785 (2003).
- Voets, T., Janssens, A., Prenen, J., Droogmans, G. & Nilius, B. Mg^{2+} -dependent gating and strong inward rectification of the cation channel TRPV6. *J. Gen. Physiol.* **121**, 245–260 (2003).
- Duncan, L. M. *et al.* Down-regulation of the novel gene melastatin correlates with potential for melanoma metastasis. *Cancer Res.* **58**, 1515–1520 (1998).
- Nagamine, K. *et al.* Molecular cloning of a novel putative Ca^{2+} channel protein (TRPC7) highly expressed in brain. *Genomics* **54**, 124–131 (1998).
- Perraud, A. L. *et al.* ADP-ribose gating of the calcium-permeable LTRPC2 channel revealed by Nudix motif homology. *Nature* **411**, 595–599 (2001).
- Sano, Y. *et al.* Immunocyte Ca^{2+} influx system mediated by LTRPC2. *Science* **293**, 1327–1330 (2001).
- Rafaty, L. A., Schmidt, M. T., Perraud, A. L., Scharenberg, A. M. & Denu, J. M. Analysis of O-acetyl-ADP-ribose as a target for Nudix ADP-ribose hydrolases. *J. Biol. Chem.* **277**, 47114–47122 (2002).
- Perraud, A. L., Schmitz, C. & Scharenberg, A. M. TRPM2 Ca^{2+} permeable cation channels: from gene to biological function. *Cell Calcium* **33**, 519–531 (2003).
- Hara, Y. *et al.* LTRPC2 Ca^{2+} -permeable channel activated by changes in redox status confers susceptibility to cell death. *Mol. Cell* **9**, 163–173 (2002).
- Heiner, I., Eisfeld, J. & Luckhoff, A. Role and regulation of TRP channels in neutrophil granulocytes. *Cell Calcium* **33**, 533–540 (2003).
- Grimm, C., Kraft, R., Sauerbruch, S., Schultz, G. & Harteneck, C. Molecular and functional characterization of the melastatin-related cation channel TRPM3. *J. Biol. Chem.* **278**, 21493–21501 (2003).
- Lee, N. *et al.* Expression and characterization of human transient receptor potential melastatin 3 (hTRPM3). *J. Biol. Chem.* **278**, 20890–20897 (2003).
- Launay, P. *et al.* TRPM4 is a Ca^{2+} -activated nonselective cation channel mediating cell membrane depolarization. *Cell* **109**, 397–407 (2002).
- Hofmann, T., Chubonov, V., Gudermann, T. & Montell, C. TRPM5 is a voltage-modulated and Ca^{2+} -activated monovalent selective cation channel. *Curr. Biol.* **13**, 1153–1158 (2003).
- Nilius, B. *et al.* Voltage dependence of the Ca^{2+} activated cation channel TRPM4. *J. Biol. Chem.* **33**, 30813–30820 (2003).
- Perez, C. A. *et al.* A transient receptor potential channel expressed in taste receptor cells. *Nature Neurosci.* **5**, 1169–1176 (2002).
- Runnels, L. W., Yue, L. & Clapham, D. E. TRP-PLIK, a bifunctional protein with kinase and ion channel activities. *Science* **291**, 1043–1047 (2001).
- Nadler, M. J. *et al.* LTRPC7 is a Mg-ATP-regulated divalent cation channel required for cell viability. *Nature* **411**, 590–595 (2001).
- Runnels, L. W., Yue, L. & Clapham, D. E. The TRPM7 channel is inactivated by PIP_2 hydrolysis. *Nature Cell Biol.* **4**, 329–336 (2002).
- Schmitz, C. *et al.* Regulation of vertebrate cellular Mg^{2+} homeostasis by TRPM7. *Cell* **114**, 191–200 (2003).
- Yamaguchi, H., Matsushita, M., Nairn, A. C. & Kuriyan, J. Crystal structure of the atypical protein kinase domain of a TRP channel with phosphotransferase activity. *Mol. Cell* **7**, 1047–1057 (2001).
- McKemy, D. D., Neuhauser, W. M. & Julius, D. Identification of a cold receptor reveals a general role for TRP channels in thermosensation. *Nature* **416**, 52–58 (2002).
- Peier, A. M. *et al.* A TRP channel that senses cold stimuli and menthol. *Cell* **108**, 705–715 (2002).
- Tracey, W. D. Jr, Wilson, R. I., Laurent, G. & Benzer, S. *painless*, a *Drosophila* gene essential for nociception. *Cell* **113**, 261–273 (2003).
- Story, G. M. *et al.* ANKTM1, a TRP-like channel expressed in nociceptive neurons, is activated by cold temperatures. *Cell* **112**, 819–829 (2003).
- Viswanath, V. *et al.* Opposite thermosensor in fruitfly and mouse. *Nature* **423**, 822–823 (2003).
- Hanaoka, K. *et al.* Co-assembly of polycystin-1 and -2 produces unique cation-permeable currents. *Nature* **408**, 990–994 (2000).
- Wu, G. *et al.* Somatic inactivation of Pkd2 results in polycystic kidney disease. *Cell* **93**, 177–188 (1998).
- Boulter, C. *et al.* Cardiovascular, skeletal, and renal defects in mice with a targeted disruption of the *Pkd1* gene. *Proc. Natl Acad. Sci. USA* **98**, 12174–12179 (2001).
- Nomura, H. *et al.* Identification of PKDL, a novel polycystic kidney disease 2-like gene whose murine homologue is deleted in mice with kidney and retinal defects. *J. Biol. Chem.* **273**, 25967–25973 (1998).
- Nonaka, S., Shiratori, H., Saijoh, Y. & Hamada, H. Determination of left-right patterning of the mouse embryo by artificial nodal flow. *Nature* **418**, 96–99 (2002).
- Nauli, S. M. *et al.* Polycystins 1 and 2 mediate mechanosensation in the primary cilium of kidney cells. *Nature Genet.* **33**, 129–137 (2003).
- Sun, M. *et al.* Mucopolidiosis type IV is caused by mutations in a gene encoding a novel transient receptor potential channel. *Hum. Mol. Genet.* **9**, 2471–2478 (2000).
- Hersh, B. M., Hartwig, E. & Horvitz, H. R. The *Caenorhabditis elegans* mucolipin-like gene *cup-5* is essential for viability and regulates lysosomes in multiple cell types. *Proc. Natl Acad. Sci. USA* **99**, 4355–4360 (2002).
- Di Palma, F. *et al.* Mutations in Mcoln3 associated with deafness and pigmentation defects in varitint-waddler (Va) mice. *Proc. Natl Acad. Sci. USA* **99**, 14994–14999 (2002).
- Hilgemann, D. W., Feng, S. & Nasuhoglu, C. The complex and intriguing lives of PIP_2 with ion channels and transporters. *Sci. STKE* **2001**, RE19 (2001).
- Hardie, R. C. Regulation of TRP channels via lipid second messengers. *Annu. Rev. Physiol.* **65**, 735–759 (2003).
- Prescott, E. D. & Julius, D. A modular PIP_2 binding site as a determinant of capsaicin receptor sensitivity. *Science* **300**, 1284–1288 (2003).
- McLaughlin, S., Wang, J., Gambhir, A. & Murray, D. PIP_2 and proteins: interactions, organization, and information flow. *Annu. Rev. Biophys. Biomol. Struct.* **31**, 151–175 (2002).
- Putney, J. W. Jr & McKay, R. R. Capacitative calcium entry channels. *Bioessays* **21**, 38–46 (1999).
- Putney, J. W. Jr TRP, inositol 1,4,5-trisphosphate receptors, and capacitative calcium entry. *Proc. Natl Acad. Sci. USA* **96**, 14669–14671 (1999).
- Prakriya, M. & Lewis, R. S. CRAC channels: activation, permeation, and the search for a molecular identity. *Cell Calcium* **33**, 311–321 (2003).
- Cui, J., Bian, J. S., Kagan, A. & McDonald, T. V. CaT1 contributes to the stores-operated calcium current in Jurkat T-lymphocytes. *J. Biol. Chem.* **277**, 47175–47183 (2002).
- Voets, T. *et al.* CaT1 and the calcium release-activated calcium channel manifest distinct pore properties. *J. Biol. Chem.* **276**, 47767–47770 (2001).
- Vaca, L. & Sampieri, A. Calmodulin modulates the delay period between release of calcium from internal stores and activation of calcium influx via endogenous TRP1 channels. *J. Biol. Chem.* **277**, 42178–42187 (2002).
- Patterson, R. L., van Rossum, D. B. & Gill, D. L. Store-operated Ca^{2+} entry: evidence for a secretion-like coupling model. *Cell* **98**, 487–499 (1999).
- Yao, Y., Ferrer-Montiel, A. V., Montal, M. & Tsien, R. Y. Activation of store-operated Ca^{2+} current in *Xenopus* oocytes requires SNAP-25 but not a diffusible messenger. *Cell* **98**, 475–485 (1999).
- Goel, M., Sinkins, W. G. & Schilling, W. P. Selective association of TRPC channel subunits in rat brain synaptosomes. *J. Biol. Chem.* **277**, 48303–48310 (2002).
- Xu, X. Z. & Sternberg, P. W. A *C. elegans* sperm TRP protein required for sperm-egg interactions during fertilization. *Cell* **114**, 285–297 (2003).
- Block, S. M. Biophysical principles of sensory transduction. *Soc. Gen. Physiol. Ser.* **47**, 1–17 (1992).
- Sukharev, S. I., Sigurdson, W. J., Kung, C. & Sachs, F. Energetic and spatial parameters for gating of the bacterial large conductance mechanosensitive channel, MscL. *J. Gen. Physiol.* **113**, 525–540 (1999).
- Howard, J., Roberts, W. M. & Hudspeth, A. J. Mechanoelectrical transduction by hair cells. *Annu. Rev. Biophys. Biophys. Chem.* **17**, 99–124 (1988).
- Kim, J. *et al.* A TRPV family ion channel required for hearing in *Drosophila*. *Nature* **424**, 81–84 (2003).
- Yellen, G. The voltage-gated potassium channels and their relatives. *Nature* **419**, 35–42 (2002).
- Hardie, R. C. *et al.* Molecular basis of amplification in *Drosophila* phototransduction: roles for G protein, phospholipase C, and diacylglycerol kinase. *Neuron* **36**, 689–701 (2002).

Acknowledgements I thank the members of the Clapham laboratory (especially H. Xu), S. McLaughlin, R. Lewis, D. Corey and C. Miller for helpful discussions, and S. Kaczmarek for artwork.

Competing interests statement The author declares that he has no competing financial interests.

Correspondence and requests for materials should be addressed to the author (dclapham@enders.tch.harvard.edu).

Structural snapshots of the mechanism and inhibition of a guanine nucleotide exchange factor

Louis Renault, Bernard Guibert & Jacqueline Cherfils

Laboratoire d'Enzymologie et Biochimie Structurales, CNRS UPR 9063, Avenue de la Terrasse, 91198 Gif sur Yvette, France

Small GTP-binding (G) proteins are activated by GDP/GTP nucleotide exchange stimulated by guanine nucleotide exchange factors (GEFs). Nucleotide dissociation from small G protein–GEF complexes involves transient GDP-bound intermediates whose structures have never been described. In the case of Arf proteins, small G proteins that regulate membrane traffic in eukaryotic cells, such intermediates can be trapped either by the natural inhibitor brefeldin A or by charge reversal at the catalytic glutamate of the Sec7 domain of their GEFs. Here we report the crystal structures of these intermediates that show that membrane recruitment of Arf and nucleotide dissociation are separate reactions stimulated by Sec7. The reactions proceed through sequential rotations of the Arf-GDP core towards the Sec7 catalytic site, and are blocked by interfacial binding of brefeldin A and unproductive stabilization of GDP by charge reversal. The structural characteristics of the reaction and its modes of inhibition reveal unexplored ways in which to inhibit the activation of small G proteins.

GEFs, which activate small G proteins by stimulating GDP/GTP exchange, are emerging as pivotal regulators in collecting the signals that lead to and determine the outcome of activation. Thus they are promising targets for interrupting signalling networks regulated by G proteins in pathological contexts¹. However, only one natural inhibitor (brefeldin A (BFA)^{2–4}) and a few examples of RNA⁵ and peptide⁶ aptamer inhibitors have been described so far. Strategies to develop inhibitors are potentially hampered by the complex multi-step mechanism of the GEFs, which involves transient G protein–GDP–GEF intermediates that isomerize to a high-affinity, nucleotide-free G protein–GEF complex^{7–9}. Structures of nucleotide-free complexes have been obtained for representative families of G protein–GEF systems^{10–13}, but low-affinity, GDP-bound intermediates have eluded structural characterization so far. Thus, how GEFs reconcile the conflicting requirements of recognizing their targets in the presence of GDP and converting them to a nucleotide-free form where the GDP-binding site is compromised remains an unanswered issue.

G proteins of the Arf family, which are major regulators of membrane dynamics in eukaryotic cells (reviewed in ref. 14), are unique in this regard because the GDP-bound Arf–GEF intermediates can be trapped at distinct successive stages of the reaction by the fungal macrolide BFA⁴ and a point mutation of the GEF¹⁵. Arf proteins operate by the classical GDP/GTP exchange common to all G proteins (reviewed in ref. 16), which they couple to a cytosol–membrane translocation (reviewed in ref. 17). This mechanism is based on an extended switch region encompassing the switch 1 and 2 regions (residues 38–52 and 69–84 in Arf1) and the intervening interswitch region comprising strands β 2 to β 3. In cytosolic Arf-GDP, the myristoylated amphipathic amino-terminal helix locks the interswitch in a retracted conformation that blocks nucleotide exchange and the binding of GTP^{18,19}. Its reversible binding to membranes releases the hasp²⁰, allowing the interswitch to break its interactions with the adjacent β -strands and shift into the helix-binding site by the ‘interswitch toggle’, thus initiating the conformation of Arf that can bind GTP¹¹. This mechanism is stimulated by GEFs carrying a catalytic Sec7 domain (reviewed in ref. 21) with an invariant glutamate (the ‘glutamic finger’), which is essential for nucleotide dissociation²² but not for the formation of

an Arf-GDP–GEF intermediate that can be trapped by mutation of the glutamic finger to lysine¹⁵. The structure of a nucleotide-free Arf–Sec7 complex revealed that the reaction has gone beyond the interswitch toggle at that stage, and that the glutamic finger excludes nucleotide binding¹¹. The exchange reaction is inhibited by BFA, a drug that acts by a rare uncompetitive mechanism with formation of an abortive Arf-GDP–BFA–Sec7 complex, which cannot proceed to nucleotide dissociation^{4,23}, and that blocks vesicular traffic in the exocytic pathway *in vivo*^{2,3}. BFA has therefore received considerable interest as a laboratory tool (reviewed in ref. 24), but also for its cytotoxic activities that include differentiation and apoptosis in cancer cell lines (ref. 25 and references therein). There are both BFA-sensitive and BFA-insensitive Arf–Sec7 systems *in vitro* and *in vivo*^{26–29}. Although the basis of the specificity of BFA for Arf isoforms is not known, residues in Sec7 domains can be mutated to convert BFA-resistant GEFs into BFA-sensitive ones and vice versa^{4,30–32}.

Here, taking advantage of point mutations that convert the BFA-insensitive Sec7 domain of the human GEF ARNO into a BFA-sensitive GEF (referred to as ARNO hereafter), we report the crystal structures of Arf1·GDP·Mg²⁺–BFA–ARNO with both full-length and N-terminally truncated Arf1, of the corresponding unbound ARNO domain, and of an abortive Arf1·GDP–ARNO complex where the glutamic finger is mutated to lysine. These structures capture the transient steps of the exchange reaction mechanism and reveal how BFA and the mutation hijack successive Arf-GDP–Sec7 intermediates. Stabilization of unproductive GEF–G protein complexes is a promising concept for the selective interruption of GEF activity in cellular networks of therapeutic interest^{4,33}, which will be discussed in the light of the structural information.

Structure of the Arf-GDP·Mg²⁺–BFA–Sec7 complex

Using human ARNO carrying the Phe190Tyr/Ala191Ser/Ser198Asp/Pro208Met point mutations, which confer high BFA sensitivity³², and bovine Arf1 (identical to human Arf1) either full length or with its N-terminal helix truncated (referred to as Arf1 and Arf1(Δ 17), respectively), the crystal structures of Arf1·GDP·Mg²⁺–BFA–ARNO and Arf1(Δ 17)·GDP·Mg²⁺–BFA–ARNO were solved at 1.9 Å and 1.8 Å, respectively (see Methods). The complexes show

that BFA targets the ArfGDP–Sec7 complex before the interswitch toggle and nucleotide dissociation but after the disruption of the extra switch 1 β -strand of ArfGDP, yielding a conformation of Arf that is distinct from both unbound ArfGDP¹⁸ and nucleotide-free Arf¹¹ (Fig. 1a). The core of Arf, excluding the extended switch region, is essentially the same as in unbound ArfGDP. The complexes with Arf1 and Arf1(Δ 17) are essentially the same, except for a distortion of the tip of the retracted interswitch into the N-terminal helix-binding site in Arf1(Δ 17) (not shown), which is artefactually made available by the truncation of the helix. Thus, the interaction of Arf1 with BFA is independent of the N-terminal helix, which explains why the BFA-inhibited complex displays a cytosol–membrane partition similar to that of ArfGDP³⁴.

The switch and interswitch of Arf form a large interface of 2,700 Å² with the hydrophobic groove and carboxy-terminal half of the Sec7 domain, creating an interfacial cavity where BFA is buried. Notably, most residues that form the interface between nucleotide-free Arf(Δ 17) and yeast Gea2 (ref. 11) already participate in the BFA-inhibited complex interface (Fig. 1a, c). Furthermore, switch 1 residues 47–53, which are inserted in the hydrophobic groove, and switch 2 residues 73–81, which sit on the Sec7 C terminus, have similar conformation and interactions in both complexes, including the hydrophobic clamps of Ile 49, Phe 51, Ile 74 and Leu 77 onto the Sec7 surface (Fig. 2). However, opening of the β 2– β 3 strands at the entrance of the interswitch, which creates the binding site for BFA (see below and Fig. 3a), results in a rotational difference of the Arf core relative to the Sec7 domain of over 20° compared with the nucleotide-free complex. This rotation keeps the GDP nucleotide remote from the glutamic finger, and accounts for the inefficiency of nucleotide exchange in the BFA-inhibited complex. The contribution of the flexibility of the Sec7 domain in binding ArfGDP and the drug was assessed by the crystal structure of unbound ARNO solved at 1.7 Å (Fig. 1a). The hydrophobic groove is open by more than 9° compared with BFA-bound ARNO, revealing that rigid-body subdomain closure is an early event rather than a movement to bring the glutamic finger near to the GDP nucleotide as previously proposed for Gea2 (ref. 35). However, there is no conformational change at the BFA-binding

site, indicating that the drug does not require an induced fit from the Sec7 domain.

Interfacial inhibition of the exchange reaction by BFA

BFA is buried in a cavity at the Arf–ARNO interface, with one-third of its surface in contact with ARNO and two-thirds with Arf. It displays close-packed hydrophobic interactions with both proteins, which combine with a few hydrogen bonds to define its orientation (Fig. 3a). The conformation of the drug is mostly unaffected by its binding to the complex compared to crystalline BFA³⁶, including the flat 13-membered lactone ring and the chirality of the C4, C7 and C14 carbons. The only difference is the puckering of the 5-membered ring, which is enclosed in an aromatic sandwich between Phe 51 and Trp 66 stacking on both sides, a site that results from the opening of the interswitch (Fig. 3b). Both chiral hydroxyls are engaged in hydrogen bonds, including O7 with the side chains of Trp 78 and Tyr 190*, and O4 with the main chain of Asp 67 (Fig. 3b; asterisk indicates a residue of the Sec7 domain). BFA also forms extended hydrophobic contacts with Trp 66 and Tyr 81 from Arf and Met 194* from ARNO (Fig. 3c). The cavity is closed by Tyr 81 in Arf1 and the 198*–208* loop in ARNO, thus addressing the question of the trajectory of the drug. Wedging of the 5-membered ring in an aromatic cul-de-sac at the interswitch entrance suggests that BFA enters the cavity by its small ring first, where it is fastened by polar interactions. Thus, the cavity should open on the opposite side, where local changes at Tyr 81 and/or the 198*–208* loop or reorientation of the Arf–Sec7 interface should modulate its access.

This model is consistent with the Sec7 domain mutations that confer BFA sensitivity (referred to hereafter by the format ‘insensitive residue/sensitive residue’ in ARNO), of which four were introduced in our ARNO construct. All sites are buried in the nucleotide-free complex¹¹ where they are not available for interaction with a ligand, but they open to the hydrophobic cavity or its first shell of interactions in the complex, where their distinct interactions with BFA account for their cumulative effects³² (Fig. 3b, c). Their mutations have only minor effects on unbound ARNO except for a slight closure of the hydrophobic groove as also found in unbound BFA-sensitive Gea2 (ref. 35), resulting from a

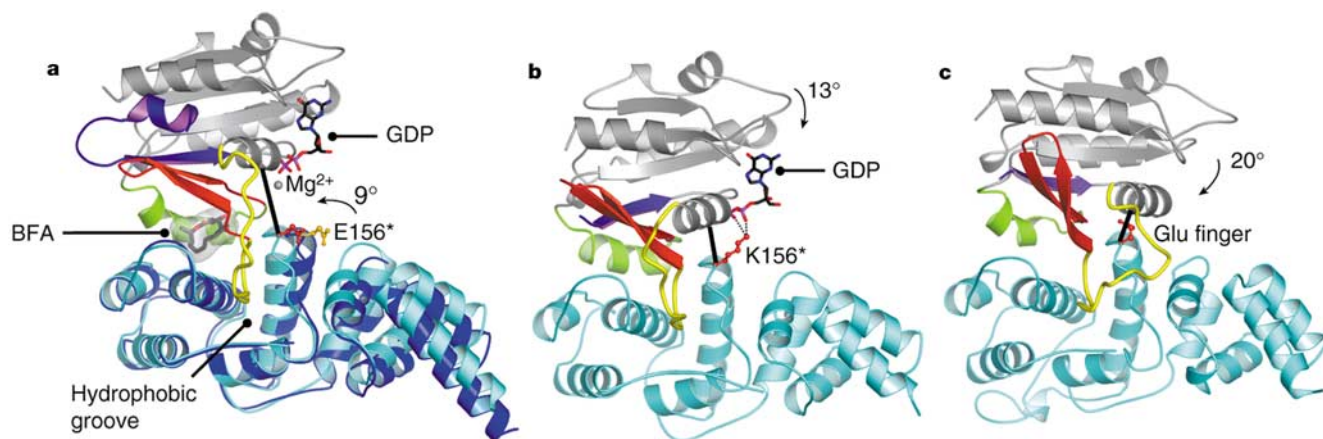


Figure 1 Snapshots of the nucleotide exchange reaction. **a**, Inhibition of the interswitch toggle by BFA in the Arf1-GDP-Mg²⁺–BFA–ARNO docking complex, with unbound ARNO in dark blue superimposed by its C terminus. Closure of the hydrophobic groove of the Sec7 domain is shown by an arrow. **b**, Inhibition of GDP dissociation by the glutamic finger mutation in the Arf1(Δ 17)-GDP–ARNO(E156K) pre-dissociation complex. The rotation of the Arf core relative to the Arf–Sec7 interface is shown by an arrow. **c**, The nucleotide-free Arf1(Δ 17)–Gea2 complex¹¹. The rotation of the Arf core towards the glutamic finger is shown by an arrow. All structures are superposed by the C terminus of their Sec7 domain, and are shown in the same orientation. The Arf core is grey, with the

N-terminal helix and strand β 1 in violet, switch 1 in yellow, interswitch in red and switch 2 in green. The Sec7 domain is in cyan, with the glutamic finger or the mutation in orange. Residues of the Sec7 domain are labelled with an asterisk. GDP and Mg²⁺ are in ball-and-stick representation and BFA is in ball-and-stick and volume representation. Approach of GDP towards the glutamic finger is shown by a black bar between Lys 30 in the P-loop and the glutamic finger. Hydrogen bonds are indicated by dotted lines. The figures are drawn using MolScript (<http://www.avatar.se/molscript>) and Pymol (<http://pymol.sourceforge.net>).

hydrogen bond of Tyr 190* across the hydrophobic groove that does not exist in wild-type ARNO²² (not shown). Two sites interact directly with BFA: Phe/Tyr 190* (ref. 4), which guides the trajectory of the drug by polar interactions, and Ile/Met 194* (refs 4, 31), which allows the lactone ring to be accommodated without steric hindrance. The Ala/Ser 191* site³² makes neither a direct nor indirect contribution to the binding of BFA, and may only have a marginal, if any, contribution. On the other hand, the Ser/Asp 198* and Pro/Met 208* sites³⁰ make only loose interactions with BFA, but they stabilize the orientation of Tyr 81 relative to the lactone and may act indirectly by modulating the access to the cavity (Fig. 3c). An indirect role of Tyr 81, as it stacks on His 80, may also explain why the His80Ala mutation stabilizes BFA in the Asp 198*/Met 208* background³². Surprisingly, a BFA-sensitizing Tyr-Ala/Asp-Met sequence is found in GBF1, whose activity on Arf5 is insensitive to BFA *in vitro*²⁷. Our structures suggest that this may result from the insertion of three residues in the loop ending at the Ser/Asp 198* and Pro/Met 208* sites, possibly obstructing the entry route followed by BFA.

Various BFA derivatives have been tested independently for either their ability to inactivate the function of Arf^{2,3} or for their cytotoxic effects on unknown targets²⁵. Configurational inversion or modification of the O4 and O7 hydroxyls consistently resulted in the loss of both properties. Our structures point to these chiral positions as key features for polar channelling of BFA into the interfacial cavity. Thus, the cytotoxic activities of BFA probably reflect its biochemical activity on Arf–Sec7 systems, supporting a previously reported correlation of the dose-dependent effects of BFA on cell growth and secretion in various cell lines³⁷. Altogether, the exquisite structural and chemical fit of BFA to the Arf–Sec7 interface accounts for its biological specificity, and for the failure of BFA derivatives—including ring-opened analogues with increased flexibility²⁵—to produce the same or enhanced cellular phenotypes.

Inhibition of GDP dissociation in Arf1(Δ17)·GDP–ARNO(E156K)

Using the Sec7 domain of ARNO carrying the mutation of the

catalytic glutamic finger (Glu 156*) to lysine in addition to the BFA-sensitizing mutations (referred to as ARNO(E156K)), the structure of the Arf1(Δ17)·GDP–ARNO(E156K) ternary complex was solved at 1.5 Å (see Methods). The structure shows that GDP is still tightly bound in an almost intact nucleotide-binding site, but the inter-switch toggle has taken place (Fig. 1b). Notably, the toggle is not mediated by a shift of the interswitch relative to the Arf–Sec7 interface (compare Fig. 3b and d), but rather by a rotation of the Arf core yielding an orientation relative to the Sec7 domain that is intermediate between those of the BFA-inhibited and the nucleotide-free complexes (Fig. 1a–c). The interface of Sec7 with the switch regions of Arf is similar in all three complexes, including the hydrophobic clamps (Fig. 2). Local optimization of the contacts at the interswitch result from residues Trp 66 and Val 68 of the β3 strand having collapsed into the hydrophobic cavity at the entrance of the interswitch, restoring the β2–β3 interactions that were disrupted in the BFA complex (Fig. 3d). The toggled interswitch obstructs the binding site of the N-terminal helix, which should therefore be displaced in full-length Arf, in agreement with the recruitment of this complex to membranes³⁴. A concomitant consequence of the rotation of the Arf core is to bring GDP close to the Sec7 catalytic site. The lysine mutation stabilizes the GDP nucleotide by interactions with its α- and β-phosphates that partially mimic those of the Mg²⁺ ion, explaining the requirement for low Mg²⁺ concentration in the formation of the complex¹⁵ (Fig. 4). This, in turn, suggests that dominant-negative effects of GEFs for Arf carrying the charge reversal mutation in cells—where they are unlikely to form abortive Arf·GDP–GEF complexes at physiological Mg²⁺ concentrations—may result from the mutants trapping cellular partners other than Arf. Thus, the mutation substitutes attractive electrostatic interactions for the repulsive interactions between the glutamic finger and the nucleotide phosphates, and acts as an obstacle to the subsequent rotation of the Arf core necessary to complete the exchange reaction. Furthermore, the lysine cannot establish the interactions carried out by the glutamic finger with the phosphate-binding P-loop, which stabilize the nucleotide-free complex¹¹.

Intermediates of the exchange reaction

Our Arf1·GDP·Mg²⁺–BFA–ARNO and Arf1(Δ17)·GDP–ARNO(E156K) structures together with the nucleotide-free Arf1(Δ17)–Gef2 complex¹¹ are consistent with a trajectory of the Arf core relative to the Sec7 domain operating on a fixed Arf–Sec7 interface via the plasticity of the extended switch region. This suggests that BFA and the mutation trap abortive complexes that are close mimics of actual docking and pre-dissociation intermediates of the exchange reaction, and can be analysed as such. Thus, this ensemble of structures provides unprecedented snapshots of the structural dynamics of an exchange reaction.

The structures reveal that the Sec7 domain catalyses the inter-switch toggle and GDP dissociation as separate sequential reactions, illustrated by the docking/pre-dissociation and pre-dissociation/nucleotide-free transitions (Fig. 1). They show that GDP is excluded from productive catalytic contacts with the Sec7 domain at the docking and pre-dissociation intermediates, and that the nucleotide approaches the glutamic finger by sequential rotations of the Arf core. The working platform for these rotations is a large Arf–Sec7 interface, which is established early and retains its characteristics throughout the reaction, including hydrophobic clamps of switch 1 and 2 onto the hydrophobic groove and the Sec7 C-terminal subdomain (Fig. 2). Formation of the interface initiates a constrained uncoupling of the Arf core relative to its extended switch region, as shown by the slight wobble of the Arf core in our three independent BFA-inhibited complexes (not shown; see Methods). The articulation between the Arf core and the Arf–Sec7 platform is orchestrated by a cluster of aromatic residues (Phe 51, Trp 66, Trp 78, Tyr 81, Phe 82 and Tyr 190*). In contrast, the Sec7 domain

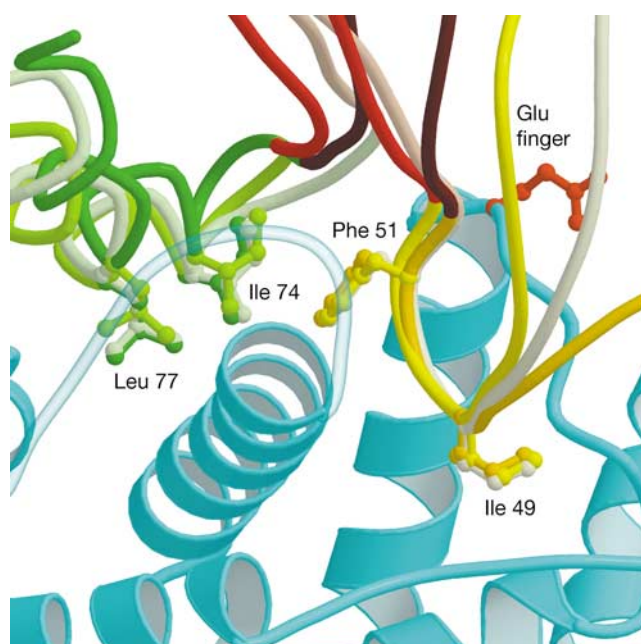


Figure 2 The Arf–Sec7 interfacial platform. Colour coding and symbols are as in Fig. 1. Arf1·GDP·Mg²⁺–BFA–ARNO, Arf1(Δ17)·GDP–ARNO(E156K) and Arf1–Gef2 (from lighter to darker shades) are superimposed based on the C terminus of their Sec7 domains. Only the Sec7 domain of Arf1·GDP·Mg²⁺–BFA–ARNO is shown for clarity. The hydrophobic clamps from switch 1 and 2 are shown.

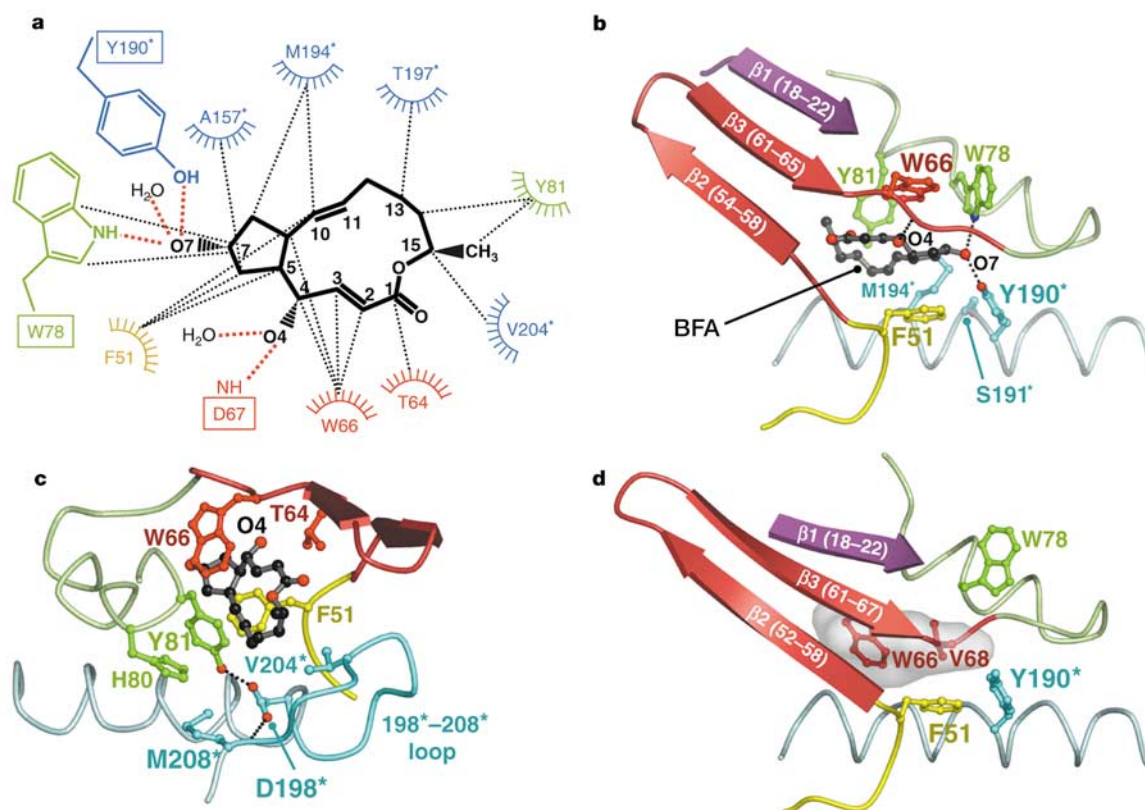


Figure 3 Interfacial inhibition of nucleotide exchange by BFA. Colour coding and symbols are as in Fig. 1. **a**, Schematic view of the interactions of BFA within the Arf–Sec7 interfacial cavity. Hydrogen bonds are indicated by red dotted lines; van der Waals contacts (cutoff of 4 Å) by black dotted lines. **b**, Close-up view of the BFA-binding site near the 5-membered ring. **c**, Close-up view of the BFA-binding site near the 13-membered

ring. The 198*–208* loop of the Sec7 domain loop is in cyan. **d**, Hydrophobic collapse of strand $\beta 3$ into the BFA-binding site in the Arf1($\Delta 17$)-GDP–ARNO(E156K) complex. The volume of BFA from the BFA-inhibited complex is superimposed in transparency. The orientation is as in Fig. 3b.

behaves as a rigid body after the early closure of its hydrophobic groove (Fig. 1a), whose probable role is to pinch off the switch 1 extra β -strand from unbound Arf-GDP by extracting Phe 51 from the Arf core.

Initiation of the Arf–Sec7 platform and subsequent rotations of the Arf core create unbalanced energetics conditions, which drive the interswitch toggle and the dissociation of GDP. First, docking of switch 1 and 2 onto the Sec7 domain disrupts the $\beta 2$ – $\beta 3$ interactions at the entrance of the interswitch, thus creating the hydrophobic cavity that is appropriated by BFA (Fig. 3b). Energetically unfavourable solvent exposure of the cavity forces the hydrophobic collapse of Trp 66 and Val 68 from $\beta 3$ (Fig. 3d), which in turn unzips from $\beta 1$. The resulting strain is relieved by the first rotation of the Arf core, which restores the $\beta 3$ – $\beta 1$ interactions with a two-residue register shift (Figs 1b and 3d), thus completing the interswitch toggle without resorting to denaturation. Concomitantly, the interswitch toggle reaction brings the nucleotide-binding site within the vicinity of the glutamic finger, such that electrostatics becomes the major driving force for subsequent structural changes (Fig. 4). It also primes the dissociation of GDP at discrete sites, including the disruption of the Lys 30–Asp 57 ionic bond, which hinders the γ -phosphate-binding site in Arf-GDP³⁸, and the Thr 31–Glu 54 hydrogen bond, which is critical for GDP affinity¹⁹ (not shown). Modelling of the glutamic finger using the lysine mutation suggests that its carboxylate, as it approaches the nucleotide, generates repulsive interactions on the β -phosphate. Notably, the ARNO(E156K) complex contains a phosphate or sulphate ion next to the β -phosphate that is engaged in multiple hydrogen bonds with a patch of positively charged residues including the

invariant Arg 162* (Fig. 4); these residues may form an electrostatic channel to help the phosphates in and out of the nucleotide-binding site. Upon nucleotide dissociation, the impaired charges of the glutamic finger and the P-loop attract each other, thereby promoting the final rotation of the Arf core and distorting the P-loop into a mimic of the phosphate oxygens¹¹.

Our structures show that either a GDP or a GTP nucleotide could be accommodated in the pre-dissociation complex. This suggests that the stable nucleotide-free complex¹¹ may be in equilibrium with an unstable nucleotide-free form resembling our ternary complex. This complex may allow non-obstructed entry of either GDP or GTP, thereby providing a structural account for the lack of nucleotide specificity of the nucleotide exchange activity⁹.

New concepts for the inhibition of GEFs

Interrupting protein networks involved in the establishment of pathological conditions is a challenge for therapeutic strategies^{33,39}. Our structures demonstrate how efficient^{2,3,24} and selective^{4,26–28} inhibition of a cellular process *in vivo* can be achieved by a drug that targets a transient protein–protein intermediate of a regulatory reaction and traps it in a non-productive conformation. Furthermore, they show that inhibition can be made highly specific by the dual interaction of the drug with both partners of the complex, which in the case of G proteins is desirable considering how GEFs outnumber their potential targets in various cellular networks. However, examples of uncompetitive interfacial inhibition of this nature are still rare⁴⁰. A few cases are illustrated by structural studies, including the immunosuppressive drug rapamycin, which traps proteins that barely interact otherwise⁴¹, and fusicoicin, a phyto-

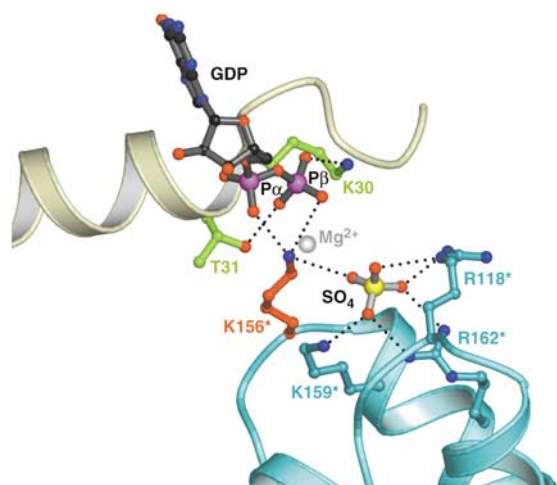


Figure 4 Stabilization of GDP by charge reversal at the glutamic finger of the Sec7 domain. The mutated lysine (K156*) that binds to the GDP phosphates and the positively charged patch of the Sec7 domain that may help nucleotides in and out of the nucleotide-binding site are shown. The expected position of the missing Mg^{2+} is shown as a grey ball. Colour coding and symbols are as in Fig. 1, residues from the phosphate-binding P-loop are in green.

toxin recently shown to over-stabilize a scaffolding complex⁴⁰. BFA extends the repertoire of uncompetitive activities by taking advantage of the structural flexibility and plasticity of its target. Arf shares these characteristics with the other members of the G-protein superfamily, thereby providing a favourable background for exploring the interfacial inhibition concept revealed by BFA.

Our observations that BFA and the mutation capture separate steps along the exchange reaction pathway suggest a dual strategy for the design or screening of uncompetitive inhibitors of G-protein activation. Binding of BFA to a retracted interswitch unique to the GDP-bound form of Arf suggests that molecules acting similarly to BFA can be explored to inhibit Arf and Arf-like proteins as these proteins share the same basic mechanism (reviewed in ref. 17), targeting for instance the BFA-insensitive Arf6 pathway but not BFA-sensitive Arf pathways. The inhibition of nucleotide exchange by the glutamic finger mutation, on the other hand, takes advantage of structural events that are probably closer to the general mechanism of GEF-catalysed G-protein activation. The mutation-inhibited complex shows that GDP keeps productive catalytic interactions at bay in the transient pre-dissociation complex, and that changes in the relative G protein–GEF orientation are necessary to yield nucleotide dissociation as reversed earlier for the related Rho/Rac G proteins⁴². The charge reversal mutation reveals that this step can be stalled by intercalating interactions that screen the nucleotide from the GEF catalytic site. This is a mechanism on which strategies to discover drugs that interrupt G protein networks of therapeutic interest could be devised^{43,44}. Together, the identification of transient exchange intermediates as potential drug targets, and the demonstration that their inhibition can be mediated by inserting obstacles at the G protein–GEF interface that stabilize abortive conformations, opens an almost unexplored avenue for the design and screening of new inhibitors of G proteins. □

Methods

Protein purification and crystallization

Wild-type bovine Arf1 (identical to human Arf1 and referred to as Arf1), Arf1 lacking the 17 N-terminal amino acids (Arf1(Δ17)), and the Sec7 domain of human ARNO (residues 50–252) carrying four BFA-sensitizing point mutations (Phe190Tyr/Ala191Ser/Ser198Asp/Pro208Met) with or without an additional Glu156Lys mutation (referred to as ARNO(E156K) and ARNO, respectively) were expressed in *Escherichia coli* and purified as described^{15,32}. All Arf–GEF complexes were prepared by incubation of a 1.4 excess of Arf proteins over Sec7 domains in the presence of inhibitor ligands and/or nucleotide, and

purified by gel filtration onto a Superdex 75 column (Amersham Biosciences). The Arf1-GDP- Mg^{2+} -BFA-ARNO complex was formed in the presence of 3 mM GDP, 4 mM $MgCl_2$ and 300 μ M BFA, and concentrated to 0.5 mM in 20 mM cacodylate pH 6.5, 50 mM ammonium formate, 50 mM ammonium fluoride, 2 mM β -mercaptoethanol with 1.25 mM GDP, 2 mM $MgCl_2$ and 500 μ M BFA. The Arf1(Δ17)-GDP- Mg^{2+} -BFA-ARNO complex was formed with 3 mM $MgCl_2$, 3 mM GDP and 300 μ M BFA and concentrated to 1.35 mM in 20 mM bicine pH 9, 100 mM K-formate, 3 mM $MgCl_2$, 3 mM β -mercaptoethanol, 2 mM DTE with 3 mM GDP and 750 μ M BFA. The Arf1(Δ17)-GDP-ARNO(E156K) complex was formed with 2 mM GDP in the presence of 8 mM EDTA and concentrated to 1.3 mM in 20 mM Tris pH 8, 25 mM K-formate, 75 mM di-ammonium hydrogen phosphate, 2 mM β -mercaptoethanol, 8 mM EDTA and 1.8 mM GDP.

All crystallizations were done by the vapour diffusion methods at approximately 16 °C by mixing equal volumes of protein and reservoir solutions in hanging drops. Crystals of unbound ARNO (concentrated to 1.1 mM in 20 mM Tris pH 8, 20 mM KCl, 2.5 mM β -mercaptoethanol) of approximately $0.1 \times 0.1 \times 0.1$ mm³ appeared within a week in 20% PEG3350, 190 mM Na-formate, 100 mM HEPES pH 8, 18 mM $MnCl_2$, were collected in the mother liquor brought to 29% PEG3350 with 6% glycerol, and were cooled in liquid ethane. Crystals of Arf1-GDP- Mg^{2+} -BFA-ARNO (approximately $0.1 \times 0.04 \times 0.04$ mm³) appeared within 1.5 months in 19% PEG3350, 175 mM ammonium fluoride, 100 mM cacodylate pH 6.5 and 17% sucrose, and were cooled directly in liquid ethane. Crystals of Arf1(Δ17)-GDP- Mg^{2+} -BFA-ARNO (approximately $0.4 \times 0.4 \times 0.04$ mm³) appeared within a week in 29% PEG3350, 40 mM ammonium chloride, 40 mM K-formate, 100 mM Tris pH 8, and were collected in the mother liquor brought to 31% PEG3350 and 5% glycerol before cooling in liquid ethane. Crystals of Arf1(Δ17)-GDP-ARNO(E156K) (approximately $0.17 \times 0.17 \times 0.06$ mm³) appeared within two weeks in 1.88 M ammonium sulphate, 100 mM MES pH 6, and were collected in the mother liquor brought to 2.3 M ammonium sulphate with 35% glycerol before cooling in liquid ethane.

Data collection, structure determination and refinement

All data collections were done at 100 K at the European Synchrotron Radiation Facility (ESRF). The data set for ARNO was merged from two incomplete data sets from different crystals. Data processing and scaling were carried out using XDS and XSCALE (<http://www.mpimf-heidelberg.mpg.de/cgi-bin/wrap/kabsch/xds>). Initial phases were obtained by molecular replacement with AmoRe or MOLREP in the CCP4 suite (<http://www.ccp4.ac.uk/>). Search models were unbound wild-type ARNO-Sec7 (ref. 22) for unbound ARNO, and Arf1(Δ17) and yeast Gea2-Sec7 from the nucleotide-free complex¹¹ for the Arf1(Δ17)-GDP- Mg^{2+} -BFA-ARNO complex. The partially refined Arf1(Δ17)-GDP- Mg^{2+} -BFA-ARNO complex was then used to solve the Arf1-GDP- Mg^{2+} -BFA-ARNO structure, and ARNO from the Arf1(Δ17)-GDP- Mg^{2+} -BFA-ARNO complex and Arf1(Δ17) from the nucleotide-free complex were used to solve the Arf1(Δ17)-GDP-ARNO(E156K) complex. Model building was done with the program O⁴⁵ and maximum likelihood refinement with bulk solvent correction was done with CNS1.1 (<http://cns.csb.yale.edu/v1.1>) and with REFMAC5 (CCP4 suite) using TLS rigid-body modelling of anisotropy. All crystals contain one complex in the asymmetric unit, except Arf1-GDP- Mg^{2+} -BFA-ARNO, which contains two complexes. Final refined models contain residues 51–246 for unbound ARNO ($R_{\text{cryst}} = 20.2\%$, $R_{\text{free}} = 23.2\%$, resolution range 33–1.7 Å); 2–180 for Arf1 and 56–246/249 for ARNO in the Arf1-GDP- Mg^{2+} -BFA-ARNO complex ($R_{\text{cryst}} = 18\%$, $R_{\text{free}} = 22.2\%$, resolution range 25–1.86 Å); 18–38 and 45–178 for Arf1(Δ17) and 57–249 for ARNO in the Arf1(Δ17)-GDP- Mg^{2+} -BFA-ARNO complex ($R_{\text{cryst}} = 18.3\%$, $R_{\text{free}} = 22.1\%$, resolution range 30–1.8 Å); and 18–177 for Arf1(Δ17) and 62–248 for ARNO(E156K) in the Arf1(Δ17)-GDP-ARNO(E156K) complex ($R_{\text{cryst}} = 16.1\%$, $R_{\text{free}} = 17.3\%$, resolution range 30–1.46 Å). Nucleotides, Mg^{2+} and BFA have clear electron densities and full occupancies in the refined structures. In Arf1-GDP- Mg^{2+} -BFA-ARNO the guanine nucleotide is guanosine 3'-monophosphate 5'-diphosphate resulting from the overexpression of Arf1 in *E. coli*, and the complex also contains a tightly bound sulphate or phosphate ion. Crystallographic statistics are summarized in Supplementary Table 1.

Received 18 September; accepted 6 November 2003; doi:10.1038/nature02197.

- Malliri, A. *et al.* Mice deficient in the Rac activator Tiam1 are resistant to Ras-induced skin tumours. *Nature* **417**, 867–871 (2002).
- Donaldson, J. G., Finazzi, D. & Klausner, R. D. Brefeldin A inhibits Golgi membrane-catalysed exchange of guanine nucleotide onto ARF protein. *Nature* **360**, 350–352 (1992).
- Helms, J. B. & Rothman, J. E. Inhibition by brefeldin A of a Golgi membrane enzyme that catalyses exchange of guanine nucleotide bound to ARF. *Nature* **360**, 352–354 (1992).
- Peyroche, A. *et al.* Brefeldin A acts to stabilize an abortive ARF-GDP-Sec7 domain protein complex: involvement of specific residues of the Sec7 domain. *Mol. Cell* **3**, 275–285 (1999).
- Mayer, G. *et al.* Controlling small guanine-nucleotide-exchange factor function through cytoplasmic RNA intramers. *Proc. Natl Acad. Sci. USA* **98**, 4961–4965 (2001).
- Schmidt, S., Diriong, S., Mery, J., Fabbriozzi, E. & Debant, A. Identification of the first Rho-GEF inhibitor, TRIPα, which targets the RhoA-specific GEF domain of Trio. *FEBS Lett.* **523**, 35–42 (2002).
- Klebe, C., Prinz, H., Wittinghofer, A. & Goody, R. S. The kinetic mechanism of Ran-nucleotide exchange catalyzed by RCC1. *Biochemistry* **34**, 12543–12552 (1995).
- Lenzen, C., Cool, R. H., Prinz, H., Kuhlmann, J. & Wittinghofer, A. Kinetic analysis by fluorescence of the interaction between Ras and the catalytic domain of the guanine nucleotide exchange factor Cdc25Mm. *Biochemistry* **37**, 7420–7430 (1998).
- Goody, R. S. & Hofmann-Goody, W. Exchange factors, effectors, GAPs and motor proteins: common thermodynamic and kinetic principles for different functions. *Eur. Biophys. J.* **31**, 268–274 (2002).
- Boriack-Sjodin, P. A., Margarit, S. M., Bar-Sagi, D. & Kuriyan, J. The structural basis of the activation of Ras by Sos. *Nature* **394**, 337–343 (1998).
- Goldberg, J. Structural basis for activation of ARF GTPase: mechanisms of guanine nucleotide exchange and GTP-myristoyl switching. *Cell* **95**, 237–248 (1998).
- Worthylake, D. K., Rossman, K. L. & Sondek, J. Crystal structure of Rac1 in complex with the guanine

- nucleotide exchange region of Tiam1. *Nature* **408**, 682–688 (2000).
13. Renault, L., Kuhlmann, J., Henkel, A. & Wittinghofer, A. Structural basis for guanine nucleotide exchange on Ran by the regulator of chromosome condensation (RCC1). *Cell* **105**, 245–255 (2001).
14. Nie, Z., Hirsch, D. S. & Randazzo, P. A. Arf and its many interactors. *Curr. Opin. Cell Biol.* **15**, 396–404 (2003).
15. Beraud-Dufour, S. *et al.* A glutamic finger in the guanine nucleotide exchange factor ARNO displaces Mg^{2+} and the β -phosphate to destabilize GDP on ARF1. *EMBO J.* **17**, 3651–3659 (1998).
16. Vetter, I. R. & Wittinghofer, A. The guanine nucleotide-binding switch in three dimensions. *Science* **294**, 1299–1304 (2001).
17. Pasqualato, S., Renault, L. & Cherfils, J. Arf, Arl, Arp and Sar proteins: a family of GTP-binding proteins with a structural device for 'front-back' communication. *EMBO Rep.* **3**, 1035–1041 (2002).
18. Amor, J. C., Harrison, D. H., Kahn, R. A. & Ringe, D. Structure of the human ADP-ribosylation factor 1 complexed with GDP. *Nature* **372**, 704–708 (1994).
19. Menetrey, J., Macia, E., Pasqualato, S., Franco, M. & Cherfils, J. Structure of Arf6-GDP suggests a basis for guanine nucleotide exchange factors specificity. *Nature Struct. Biol.* **7**, 466–469 (2000).
20. Antonny, B., Beraud-Dufour, S., Chardin, P. & Chabre, M. N-terminal hydrophobic residues of the G-protein ADP-ribosylation factor-1 insert into membrane phospholipids upon GDP to GTP exchange. *Biochemistry* **36**, 4675–4684 (1997).
21. Jackson, C. L. & Casanova, J. E. Turning on ARF: the Sec7 family of guanine-nucleotide-exchange factors. *Trends Cell Biol.* **10**, 60–67 (2000).
22. Cherfils, J. *et al.* Structure of the Sec7 domain of the Arf exchange factor ARNO. *Nature* **392**, 101–105 (1998).
23. Mansour, S. J. *et al.* p200 ARF-GEPI1: a Golgi-localized guanine nucleotide exchange protein whose Sec7 domain is targeted by the drug brefeldin A. *Proc. Natl Acad. Sci. USA* **96**, 7968–7973 (1999).
24. Jackson, C. L. Brefeldin A revealing the fundamental principles governing membrane dynamics and protein transport. *Subcell. Biochem.* **34**, 233–272 (2000).
25. Zhu, J. W. *et al.* Elucidation of strict structural requirements of brefeldin A as an inducer of differentiation and apoptosis. *Bioorg. Med. Chem.* **8**, 455–463 (2000).
26. Chardin, P. *et al.* A human exchange factor for ARF contains Sec7- and pleckstrin-homology domains. *Nature* **384**, 481–484 (1996).
27. Claude, A. *et al.* GBF1: A novel Golgi-associated BFA-resistant guanine nucleotide exchange factor that displays specificity for ADP-ribosylation factor 5. *J. Cell Biol.* **146**, 71–84 (1999).
28. Donaldson, J. G. & Jackson, C. L. Regulators and effectors of the ARF GTPases. *Curr. Opin. Cell Biol.* **12**, 475–482 (2000).
29. Donaldson, J. G. Multiple roles for Arf6: Sorting, structuring, and signaling at the plasma membrane. *J. Biol. Chem.* **278**, 41573–41576 (2003).
30. Sata, M., Moss, J. & Vaughan, M. Structural basis for the inhibitory effect of brefeldin A on guanine nucleotide-exchange proteins for ADP-ribosylation factors. *Proc. Natl Acad. Sci. USA* **96**, 2752–2757 (1999).
31. Baumgartner, F., Wiek, S., Paprotka, K., Zauner, S. & Lingelbach, K. A point mutation in an unusual Sec7 domain is linked to brefeldin A resistance in a *Plasmodium falciparum* line generated by drug selection. *Mol. Microbiol.* **41**, 1151–1158 (2001).
32. Robineau, S., Chabre, M. & Antonny, B. Binding site of brefeldin A at the interface between the small G protein ADP-ribosylation factor 1 (ARF1) and the nucleotide-exchange factor Sec7 domain. *Proc. Natl Acad. Sci. USA* **97**, 9913–9918 (2000).
33. Chardin, P. & McCormick, F. Brefeldin A: the advantage of being uncompetitive. *Cell* **97**, 153–155 (1999).
34. Beraud-Dufour, S., Paris, S., Chabre, M. & Antonny, B. Dual interaction of ADP ribosylation factor 1 with Sec7 domain and with lipid membranes during catalysis of guanine nucleotide exchange. *J. Biol. Chem.* **274**, 37629–37636 (1999).
35. Renault, L., Christova, P., Guibert, B., Pasqualato, S. & Cherfils, J. Mechanism of domain closure of Sec7 domains and role in BFA sensitivity. *Biochemistry* **41**, 3605–3612 (2002).
36. Weber, H. P., Hauser, D. & Sigg, H. P. Structure of brefeldin A. *Helv. Chim. Acta* **54**, 2763–2766 (1971).
37. Torii, S. *et al.* Cytotoxicity of brefeldin A correlates with its inhibitory effect on membrane binding of COP coat proteins. *J. Biol. Chem.* **270**, 11574–11580 (1995).
38. Pasqualato, S., Renault, L. & Cherfils, J. In *The ARF Book* (ed. Kahn, R. A.) Ch. 3 (Kluwer Academic, Dordrecht, in the press).
39. McCormick, F. Small-molecule inhibitors of cell signaling. *Curr. Opin. Biotechnol.* **11**, 593–597 (2000).
40. Wurtele, M., Jelich-Ottmann, C., Wittinghofer, A. & Oecking, C. Structural view of a fungal toxin acting on a 14-3-3 regulatory complex. *EMBO J.* **22**, 987–994 (2003).
41. Choi, J., Chen, J., Schreiber, S. L. & Clardy, J. Structure of the FKBP12-rapamycin complex interacting with the binding domain of human FRAP. *Science* **273**, 239–242 (1996).
42. Cherfils, J. Structural mimicry of DH domains by Arfaptin suggests a model for the recognition of Rac-GDP by its guanine nucleotide exchange factors. *FEBS Lett.* **507**, 280–284 (2001).
43. Cox, A. D. & Der, C. J. Ras family signaling: therapeutic targeting. *Cancer Biol. Ther.* **1**, 599–606 (2002).
44. Boettner, B. & Van Aelst, L. The role of Rho GTPases in disease development. *Gene* **286**, 155–174 (2002).
45. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47**, 110–119 (1991).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank the ESRF for provision of synchrotron radiation facilities and the staff of beamlines BM30A, ID14-EH2 and ID14-EH3 for assistance. This work was supported by grants from the Association pour la Recherche contre le Cancer, the CNRS, the Ministère de la Recherche and the Human Frontier Science Program. We thank B. Antonny and co-workers at the IPMC (CNRS) for the gift of the plasmids used in this work and for discussions. The Arf1(Δ 17)–Gea2–Sec7 coordinates are a gift from J. Goldberg.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to J.C. (cherfils@lebs.cnrs-gif.fr). Coordinates and structure factors have been deposited in the Protein Data Bank under accession numbers 1R8M and RCSB020572 for unbound ARNO, 1R8Q and RCSB020576 for Arf1-GDP-Mg²⁺–BFA–ARNO, 1R8R and RCSB020577 for Arf1(Δ 17)-GDP-Mg²⁺–BFA–ARNO and 1R8S and RCSB020578 for Arf1(Δ 17)-GDP–ARNO(E156K).

An increased estimate of the merger rate of double neutron stars from observations of a highly relativistic system

M. Burgay¹, N. D'Amico^{2,3}, A. Possenti^{3,4}, R. N. Manchester⁵, A. G. Lyne⁶, B. C. Joshi^{6,7}, M. A. McLaughlin⁶, M. Kramer⁶, J. M. Sarkissian⁵, F. Camilo⁸, V. Kalogera⁹, C. Kim⁹ & D. R. Lorimer⁶

¹Università degli Studi di Bologna, Dipartimento di Astronomia, via Ranzani 1, 40127, Bologna, Italy

²Università degli Studi di Cagliari, Dipartimento di Fisica, SP Monserrato-Sestu km 0.7, 09042 Monserrato, Italy

³INAF—Osservatorio Astronomico di Cagliari, Loc. Poggio dei Pini, Strada 54, 09012 Capoterra, Italy

⁴INAF—Osservatorio Astronomico di Bologna, via Ranzani 1, 40127, Bologna, Italy

⁵Australia Telescope National Facility, CSIRO, PO Box 76, Epping, New South Wales 2121, Australia

⁶University of Manchester, Jodrell Bank Observatory, Macclesfield, Cheshire, SK11 9DL, UK

⁷National Center for Radio Astrophysics, PO Bag 3, Ganeshkhind, Pune 411007, India

⁸Columbia Astrophysics Laboratory, Columbia University, 550 West 120th Street, New York, New York 10027, USA

⁹Northwestern University, Department of Physics and Astronomy, Evanston, Illinois 60208, USA

The merger¹ of close binary systems containing two neutron stars should produce a burst of gravitational waves, as predicted by the theory of general relativity². A reliable estimate of the double-neutron-star merger rate in the Galaxy is crucial in order to predict whether current gravity wave detectors will be successful in detecting such bursts. Present estimates of this rate are rather low^{3–7}, because we know of only a few double-neutron-star binaries with merger times less than the age of the Universe. Here we report the discovery of a 22-ms pulsar, PSR J0737–3039, which is a member of a highly relativistic double-neutron-star binary with an orbital period of 2.4 hours. This system will merge in about 85 Myr, a time much shorter than for any other known neutron-star binary. Together with the relatively low radio luminosity of PSR J0737–3039, this timescale implies an order-of-magnitude increase in the predicted merger rate for double-neutron-star systems in our Galaxy (and in the rest of the Universe).

PSR J0737–3039 was discovered during a pulsar search carried out using a multibeam receiver⁸ on the Parkes 64-m radio telescope in New South Wales, Australia. The original detection showed a large change in apparent pulsar period during the 4-min observation time, suggesting that the pulsar is a member of a tight binary system. Follow-up observations undertaken at Parkes consisting of continuous ~5-h observations showed that the orbit has a very short period (2.4 h) and a significant eccentricity (0.088). The derived orbital parameters implied that the system is relatively massive, probably consisting of two neutron stars, and predicted a huge rate of periastron advance $\dot{\omega}$ due to effects of general relativity. Indeed, after only a few days of pulse-timing observations we were able to detect a significant value of $\dot{\omega}$.

Interferometric observations made using the Australia Telescope Compact Array (ATCA) in the 20-cm band gave an improved position and flux density for the pulsar. Knowledge of the pulsar position with subarcsecond precision allowed determination of the rotational period derivative, $\dot{P}$, and other parameters from the available data span. Table 1 reports results derived from a coherent phase fit to data taken over about five months. The measured value

of $\dot{\omega} = 16.88^\circ \text{yr}^{-1}$ is about four times that of PSR B1913+16 (ref. 9), previously the highest-known. If the observed $\dot{\omega}$ is entirely due to general relativity, it indicates a total system mass $M = 2.58 \pm 0.02 M_\odot$, where $M_\odot$ is the mass of the Sun.

Figure 1 shows the constraints on the masses of the pulsar and its companion resulting from the observations so far and the mean pulse profile as an inset. The shaded region indicates values that are ruled out by the mass function M_f and the observed $\dot{\omega}$ constrains the system to lie between the two diagonal lines. Together, these constraints imply that the pulsar mass m_p is less than $1.35 M_\odot$ and that the companion mass m_c is greater than $1.24 M_\odot$. The derived upper limit on m_p is consistent with the mass distribution of pulsars¹⁰ in double-neutron-star systems. The minimum allowed companion mass implies that the companion is very likely also to be a neutron star.

This is supported by the observed significant eccentricity of the orbit, indicative of sudden and/or asymmetric mass loss from the system in the explosion when the companion collapsed¹¹. Significant classical contributions to $\dot{\omega}$, either from tidal deformation¹² or rotation-induced quadrupole moment^{13,14} in the companion—effects which might be present if it were not a neutron star—are therefore unlikely. In fact, $\dot{\omega}_{\text{GR}} \ll \dot{\omega}$ would imply an unacceptably low value for the pulsar mass. Hence this system probably contains two neutron stars with relatively low masses in the range 1.24–1.35 $M_\odot$, and the binary orbit is probably nearly edge-on (inclination, $i \approx 90^\circ$). For $i = 68^\circ$, the implied mass of the pulsar is $1.24 M_\odot$ and that of the companion is $1.35 M_\odot$.

A precise determination of the masses of both stars, and then of the orbital inclination, will be obtained when the gravitational redshift⁹, γ , is measured. The expected minimum value (assuming an edge-on orbit) of this parameter is 383 μs . However, its orbital dependence is covariant with $a \sin i$ (Table 1) until there has been significant periastron precession. Whereas we currently get a formal value of $\gamma = 400 \pm 100 \mu\text{s}$ from a fit of the timing data, a useful measurement of γ should be possible within a year.

Table 1 Measured and derived parameters for PSR J0737–3039

Right ascension (J2000)	07 h 37 min 51.28(2) s
Declination (J2000)	−30° 39′ 40.3(4)″
Galactic longitude	245.24°
Galactic latitude	−4.50°
Dispersion measure (pc cm ^{−3})	48.9(2)
Flux density at 1,400 MHz (mJy)	6.9(6)
Flux density at 430 MHz (mJy)	100(20)
Period, P (ms)	22.69937854062(6)
Period derivative, $\dot{P}$ (s s ^{−1})	$2.3(6) \times 10^{-18}$
Epoch (MJD)	52800.0
Orbital period, P_b (days)	0.102251561(8)
Projected semi-major axis, $a \sin i$ (light s)	1.415176(5)
Eccentricity	0.087793(8)
Longitude of periastron, ω	68.743(9)°
Epoch of periastron, T_0 (MJD)	52760.807391(3)
Advance of periastron, $\dot{\omega}$ (° yr ^{−1})	16.88(10)
Post-fit residual (μs)	33
Mass function, M_f ($M_\odot$)	0.291054(8)
Total system mass, M ($M_\odot$)	2.58(2)
Characteristic age, τ_c (Myr)	160(50)
Age since birth, τ_b (Myr)	100(50)
Surface dipole magnetic field strength, B (G)	$7.3(9) \times 10^9$
Expected gravitational redshift parameter, γ (μs)	>383
Expected orbital period derivative (s s ^{−1})	-1.24×10^{-12}

The pulsar position is obtained from observations made with the ATCA used in pulsar-gating mode. Uncertainties are given in parentheses and refer to the last quoted digit. The flux at 1,400 MHz is derived from the ATCA observations, while the flux at 430 MHz is obtained from Parkes observations with the flux value being derived from the observed signal-to-noise ratio and the estimated receiver sensitivity. Best-fit timing parameters are derived from arrival-time data by minimizing the χ^2 function using the program TEMPO (see <http://pulsar.princeton.edu/tempo>) with the relativistic binary model of ref. 27. For all the timing parameters except P_b , $\dot{\omega}$ and $\dot{P}$, uncertainties are twice the formal-fit 1 σ error. Because of the short time span, $\dot{\omega}$ is strongly covariant with other orbital parameters and the nominal 1 σ error resulting from the phase coherent fit of the pulse time-of-arrival might not represent the actual 1 σ statistical uncertainty. By producing contours of constant χ^2 values as a function of P_b and $\dot{\omega}$, we have verified that an uncertainty of the order of four times the nominal 1 σ error represents the uncertainty range for these parameters. The uncertainty in $\dot{P}$ is due to the residual covariance with the position terms resulting from the available uncertainty in the celestial coordinates obtained at the ATCA. The characteristic age τ_c and the surface dipole magnetic field strength B are derived from the observed P and $\dot{P}$ using the relations²⁸ $\tau_c = 0.5P/\dot{P}$ and $B = 3.2 \times 10^{19} (PP\dot{P})^{1/2}$. The age since birth, τ_b , is the time since the binary pulsar left the spin-up line⁴.

The parameters of the PSR J0737–3039 binary system, and the high timing precision made possible by the narrow pulses of this pulsar, promise to make this system an excellent ‘laboratory’ for the investigation of relativistic astrophysics, in particular of all the effects that are enhanced by a short orbital period, P_b . Geodetic spin-precession due to curvature of space-time around the pulsar is expected to have a period of 75 yr. This is four times less than the previously known shortest spin precession period and should allow detailed mapping of the pulsar emission beam¹⁵ in a few years. It may also have the undesirable effect of making the pulsar undetectable in a short time.

The secular decay of the orbital period due to gravitational wave emission should be measurable within about 15 months. In PSR B1913+16, uncertainty in the orbital period derivative due to acceleration in the Galactic potential limits the precision of this test of the predictions of general relativity^{16,17}. By comparison, PSR J0737–3039 is relatively close to the Sun and the uncertainty in Galactic acceleration will be much smaller, allowing a more precise test. Because of the probable high orbital inclination of the system, the delay due to the curvature of the space-time in the vicinity of the companion (the Shapiro delay¹⁸) should be measurable.

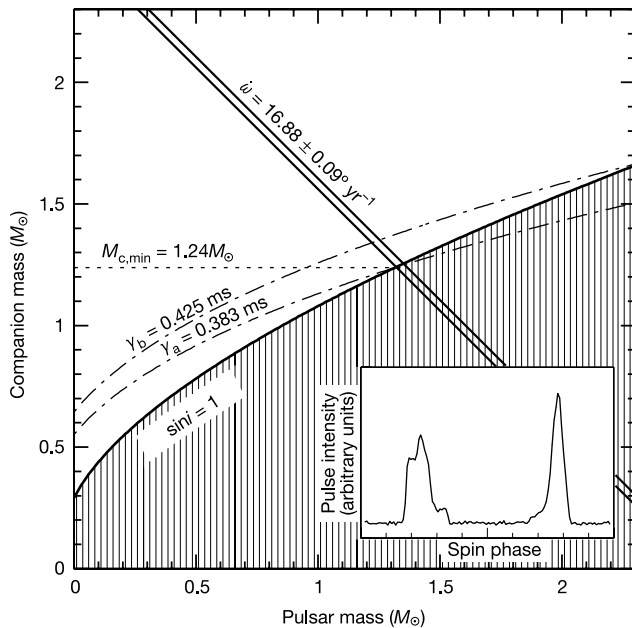


Figure 1 Constraints on the masses of the neutron star and the companion. The keplerian orbital parameters give the mass function:

$$M_f = \frac{(m_c \sin i)^3}{(m_c + m_p)^2} = \frac{4\pi^2 (a \sin i)^3}{GP_b^2} = 0.291054(8) M_\odot$$

where m_p and m_c are the masses of the pulsar and companion, P_b is the orbital period, and $a \sin i$ is the projected semi-major axis of the pulsar orbit; the estimated uncertainty in the last quoted digit is given in parentheses. This provides a constraint on the minimum m_c as a function of m_p , by setting $i = 90^\circ$, where i is the inclination of the orbit normal to the line of sight. In addition, assuming that the periastron advance rate $\dot{\omega}$ is entirely due to general relativity (that is, $\dot{\omega} = \dot{\omega}_{GR}$), its value can be combined with the other orbital parameters to determine the total system mass $M = m_c + m_p$ using the equation²⁷:

$$m_c + m_p = \frac{P_b^{5/2}}{2\pi G} \left(\frac{(1 - e^2)c^2 \dot{\omega}}{6\pi} \right)^{3/2} = 2.58(2) M_\odot$$

where e is the orbital eccentricity. In the $m_c - m_p$ plane, the shaded region indicates the mass values ruled out by the mass function M_f , and the intersection of the boundary of this region with the $\dot{\omega}$ limits (defining a diagonal strip in the plot) provides a further constraint on the masses. The dash-dotted lines represent the mass values corresponding to the relativistic parameter γ for an edge-on orbit (γ_a) and for an orbital inclination $i = 68^\circ$ (γ_b). The inset shows the mean pulse profile at 1,400 MHz from a 5-h observation. There are two pulse components separated by approximately half the 22-ms period.

The relatively small orbital eccentricity suggests that this system may have already undergone a substantial change of the orbital elements due to gravitational decay. Assuming that the evolution of orbital separation and eccentricity are entirely driven by gravitational radiation emission¹⁹, and adopting a value of 100 Myr for the time elapsed since birth⁴, τ_b (see Table 1), the orbital initially had a period of 3.3 h and an eccentricity of 0.119. Hence this system would already have experienced a $\sim 25\%$ decay of the orbital period and circularization, which would be the largest variation of these parameters in the sample of the known double-neutron-star binary systems.

Of the five double-neutron-star systems known so far²⁰, only three have orbits tight enough that the two neutron stars will merge within a Hubble time. Two of them (PSR B1913+16 and PSR B1534+12) are located in the Galactic field, while the third (PSR B2127+11C) is found on the outskirts of a globular cluster. The contribution of globular cluster systems to the Galactic merger rate is estimated to be negligible²¹. Also, recent studies⁶ have demonstrated that the current estimate of the Galactic merger rate R relies mostly on PSR B1913+16. So, in the following discussion, we start by comparing the observed properties of PSR B1913+16 and PSR J0737–3039.

PSR J0737–3039 and its companion star will merge due to the emission of gravitational waves in $\tau_m \approx 85$ Myr, a timescale that is a factor of 3.5 shorter than that for PSR B1913+16 (ref. 9). In addition, the estimated distance of PSR J0737–3039 (500–600 pc, based on the observed dispersion measure and a model for the distribution of ionized gas in the interstellar medium²²) is an order of magnitude less than that of PSR B1913+16. These properties have a substantial effect on the prediction of the rate of merging events in the Galaxy.

For a given class k of binary pulsars in the Galaxy, apart from a beaming correction factor, the merger rate R_k is calculated⁶ as $R_k \propto N_k / \tau_k$. Here τ_k is the binary pulsar lifetime defined as the sum of the time since birth⁴, τ_b , and the remaining time before coalescence, τ_m , and N_k is the scaling factor defined as the number of binaries in the Galaxy belonging to the given class. The shorter lifetime of PSR J0737–3039 ($\tau_{1913} / \tau_{0737} = (365 \text{ Myr}) / (185 \text{ Myr}) \approx 2$, where the subscript numbers refer to the pulsars), implies a doubling of the ratio R_{0737} / R_{1913} . A much more substantial increase results from the computation of the ratio of the scaling factors N_{0737} / N_{1913} . The luminosity $L_{400} \approx 30 \text{ mJy kpc}^2$ of PSR J0737–3039 is much lower than that of PSR B1913+16 ($\sim 200 \text{ mJy kpc}^2$). For a planar homogeneous distribution of pulsars in the Galaxy, the ratio $N_{0737} / N_{1913} \propto L_{1913} / L_{0737} \approx 6$. Hence we obtain $R_{0737} / R_{1913} \approx 12$. Including the moderate contribution of the longer-lived PSR B1534+12 system to the total rate^{4–7}, we obtain an increase factor for the total merger rate $(R_{0737} + R_{1913} + R_{1534}) / (R_{1913} + R_{1534})$ of about an order of magnitude. A better estimate of this factor and its uncertainty can be obtained using a bayesian statistical approach and accounting for the full luminosity function of pulsars in the Galaxy⁶.

Figure 2 displays an estimate of the probability density function for the merger rate increase factor, showing a peak value of ~ 8 and an upper limit of ~ 30 at a 95% confidence level. This result, derived in the simple assumption of a fixed pulsar luminosity and a uniform disk distribution of pulsar binaries, can be refined by including the parameters of the present survey into a simulation program modelling survey selection effects (and hence detection probability) and the Galactic population of pulsars. The Parkes high-latitude survey covers the region enclosed by Galactic latitude $|b| < 60^\circ$ and Galactic longitude $220^\circ < l < 260^\circ$ for a total of 6,456 pointings, each of duration 265 s, using a bandwidth of 288 MHz and a sampling time of 125 μ s. PSR J0737–3039 was detected with a signal-to-noise ratio of 18.7 in a standard Fourier search²³. Variation in the apparent pulsar period due to binary motion during the discovery observation resulted in a 30% reduction in the observed signal-to-noise ratio.

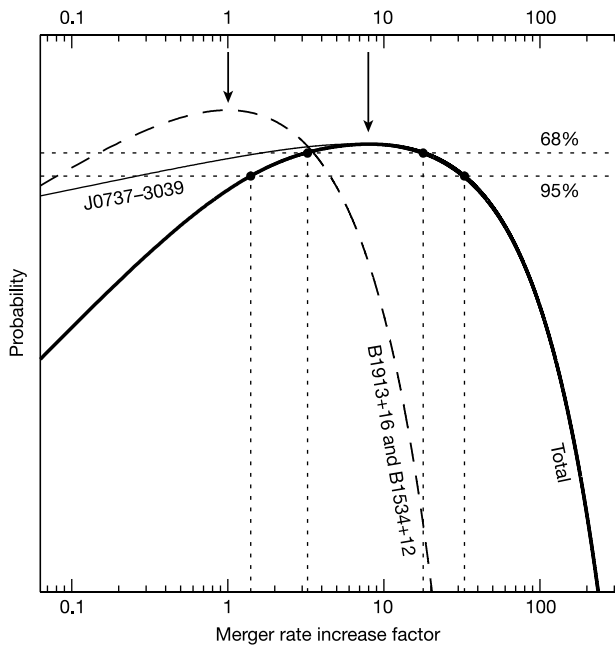


Figure 2 Probability density function for the increase in the double-neutron-star merger rate $(R_{0737} + R_{1913} + R_{1534})/(R_{1913} + R_{1534})$ resulting from the discovery of PSR J0737–3039. For a given class k of binary pulsars in the Galaxy, the probability density function $P(R_k)$ for the corresponding merger rate R_k is obtained from the relation⁶ $P(R_k) = A^2 R_k e^{-AR_k}$ where $A = \tau_k/(N_k f_k)$, N_k is a population scaling factor (see text), τ_k is the binary pulsar lifetime and f_k is the correction factor due to the beamed nature of the radio pulsar emission. The dashed line represents the reference probability density function corresponding to the merger rate $R_{\text{old}} = R_{1913} + R_{1534}$ due to PSR B1913+16 and PSR B1534+12 only⁶. The heavy solid line represents the probability density function corresponding to the new merger rate $R_{\text{new}} = R_{0737} + R_{1913} + R_{1534}$. The parameters adopted for PSR J0737–3039 are those derived in the text ($N_{0737} = 6N_{1913}$, $\tau_{0737} = 185$ Myr) and the beaming factor for PSR J0737–3039, f_{0737} , is chosen as the average of the beaming factor of the other two known merging binaries ($f_{0737} = 1.06f_{1913}$). The dotted vertical lines represent the 68% and 95% confidence level boundaries on the determination of the increase factor. The dominant role of PSR J0737–3039 in shaping the new statistics of the double-neutron-star merger rate is evident.

Extensive simulations (V.K. *et al.*, manuscript in preparation) produce results consistent with that derived in Fig. 2, and show that the peak of the merger rate increase factor resulting from the discovery of PSR J0737–3039 lies in the range 6 to 7 and is largely independent of the adopted pulsar population model. On the other hand, the actual predicted value of the Galactic merger rate and hence the detection rate by gravity wave detectors depends on the shape of the pulsar luminosity function. For the most favourable distribution model available (model 15 of ref. 6), the updated cosmic detection rate for first-generation gravity wave detectors such as VIRGO²⁴, LIGO²⁵ and GEO²⁶ can be as high as 1 every 1–2 yr at the 95% confidence level.

After a few years of operation of the gravity wave detectors, it should be possible to test these predictions directly and thus place better constraints on the cosmic population of double-neutron-star binaries. □

Received 12 August; accepted 15 October 2003; doi:10.1038/nature02124.

1. Misner, C., Thorne, K. S. & Wheeler, J. A. *Gravitation* Ch. 36 (Freeman, New York, 1973).
2. Schutz, B. GW, sources, and physics overview: Proc. 5th Edoardo Amaldi Conf. on Gravitational Waves. *Class. Quant. Grav.* (special issue) (in the press).
3. Curran, S. J. & Lorimer, D. R. Pulsar statistics. Part 3: Neutron star binaries. *Mon. Not. R. Astron. Soc.* **276**, 347–352 (1995).
4. Arzumanyan, Z., Cordes, J. M. & Wasserman, I. Pulsar spin evolution, kinematics, and the birthrate of neutron star binaries. *Astrophys. J.* **520**, 696–705 (1999).
5. Kalogera, V., Narayan, R., Spiegel, D. N. & Taylor, J. H. The coalescence rate of double neutron star systems. *Astrophys. J.* **556**, 340–356 (2001).

6. Kim, C., Kalogera, V. & Lorimer, D. R. The probability distribution of binary pulsar coalescence rates. I. Double neutron star systems in the galactic field. *Astrophys. J.* **584**, 985–995 (2003).
7. van den Heuvel, E. P. J. & Lorimer, D. R. On the Galactic and cosmic merger rate of double neutron stars. *Mon. Not. R. Astron. Soc.* **283**, L37–L39 (1996).
8. Staveley-Smith, L. *et al.* The Parkes 21 cm multibeam receiver. *Publ. Astron. Soc. Aust.* **13**, 243–248 (1996).
9. Taylor, J. H., Fowler, L. A. & McCulloch, P. M. Measurements of general relativistic effects in the binary pulsar PSR 1913+16. *Nature* **277**, 437–440 (1979).
10. Thorsett, S. E. & Chakrabarty, D. Neutron star mass measurements. I. Radio pulsars. *Astrophys. J.* **512**, 288–299 (1999).
11. Srinivasan, G. & van den Heuvel, E. P. J. Some constraints on the evolutionary history of the binary pulsar PSR 1913+16. *Astron. Astrophys.* **108**, 143–147 (1982).
12. Roberts, D. H., Masters, A. R. & Arnett, W. D. Determining the stellar masses in the binary system containing the pulsar PSR 1913+16—Is the companion a helium main-sequence star? *Astrophys. J.* **203**, 196–201 (1976).
13. Smarr, L. L. & Blandford, R. The binary pulsar—Physical processes, possible companions, and evolutionary histories. *Astrophys. J.* **207**, 574–588 (1976).
14. Wex, N. A timing formula for main-sequence star binary pulsars. *Mon. Not. R. Astron. Soc.* **298**, 66–77 (1998).
15. Weisberg, J. M. & Taylor, J. H. General relativistic geodetic spin precession in binary pulsar B1913+16: Mapping the emission beam in two dimensions. *Astrophys. J.* **576**, 942–949 (2002).
16. Damour, T. & Taylor, J. H. On the orbital period change of the binary pulsar PSR 1913+16. *Astrophys. J.* **366**, 501–511 (1991).
17. Taylor, J. H. Binary pulsars and relativistic gravity. *Rev. Mod. Phys.* **66**, 711–719 (1994).
18. Shapiro, I. I. Fourth test of general relativity. *Phys. Rev. Lett.* **13**, 789–791 (1964).
19. Peters, P. C. & Mathews, J. Gravitational radiation from point masses in a keplerian orbit. *Phys. Rev.* **131**, 435–440 (1963).
20. Taylor, J. H. Pulsar timing and relativistic gravity. *Phil. Trans. R. Soc. Lond.* **341**, 117–134 (1992).
21. Phinney, E. S. The rate of neutron star binary mergers in the universe—Minimal predictions for gravity wave detectors. *Astrophys. J.* **380**, L17–L21 (1991).
22. Taylor, J. H. & Cordes, J. M. Pulsar distances and the galactic distribution of free electrons. *Astrophys. J.* **411**, 674–684 (1993).
23. Manchester, R. N. *et al.* The Parkes multi-beam pulsar survey—I. Observing and data analysis systems, discovery and timing of 100 pulsars. *Mon. Not. R. Astron. Soc.* **328**, 17–35 (2001).
24. Bradaschia, C., *et al.* in *Gravitational Astronomy: Instrument Design and Astrophysical Prospects* (eds McClelland, D. E. & Bachor, H. A.) 110–115 (Elizabeth and Frederick White Research Conference Proceedings, World Scientific, Singapore, 1991).
25. Abramovici, A. *et al.* LIGO—The Laser Interferometer Gravitational-Wave Observatory. *Science* **256**, 325–333 (1992).
26. Danzmann, K., *et al.* in *First Edoardo Amaldi Conf. on Gravitational Wave Experiments* (eds Coccia, E., Pizzella, G. & Ronga, F.) 100–111 (World Scientific, Singapore, 1995).
27. Damour, T. & Deruelle, N. General relativistic celestial mechanics of binary systems. II. The post-Newtonian timing formula. *Ann. Inst. H. Poincaré (Phys. Théor.)* **44**, 263–292 (1986).
28. Manchester, R. N. & Taylor, J. H. *Pulsars* Ch. 9 (Freeman, San Francisco, 1977).

Acknowledgements We thank J. Reynolds of the Parkes Observatory, and B. Saul of the ATCA, for prompt allocations of observing time. The Parkes Observatory and the ATCA are part of the Australia Telescope, which is funded by the Commonwealth of Australia for operation as a National Facility managed by CSIRO. M.B., N. D'A. and A.P. acknowledge financial support from the Italian Ministry of University and Research (MIUR) under the national programme 'Cofin 2001'. V.K. acknowledges partial support by a David and Lucile Packard Science and Engineering Fellowship and a NSF Gravitational Physics grant. D.R.L. is a University Research fellow funded by the Royal Society.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to N. D'A. (damico@ca.astro.it).

Continuous magnetic reconnection at Earth's magnetopause

H. U. Frey¹, T. D. Phan¹, S. A. Fuselier² & S. B. Mende¹

¹Space Sciences Laboratory, University of California, Berkeley, California 94720-7450, USA

²Lockheed Martin Advanced Technology Center, Palo Alto, California 94304, USA

The most important process that allows solar-wind plasma to cross the magnetopause and enter Earth's magnetosphere is the merging between solar-wind and terrestrial magnetic fields of opposite sense—magnetic reconnection¹. It is at present not known whether reconnection can happen in a continuous fashion or whether it is always intermittent. Solar flares² and magneto-

spheric substorms³—two phenomena believed to be initiated by reconnection—are highly burst-like occurrences, raising the possibility that the reconnection process is intrinsically intermittent, storing and releasing magnetic energy in an explosive and uncontrolled manner. Here we show that reconnection at Earth's high-latitude magnetopause is driven directly by the solar wind, and can be continuous and even quasi-steady over an extended period of time. The dayside proton auroral spot in the ionosphere—the remote signature of high-latitude magnetopause reconnection⁴—is present continuously for many hours. We infer that reconnection is not intrinsically intermittent; its steadiness depends on the way that the process is driven.

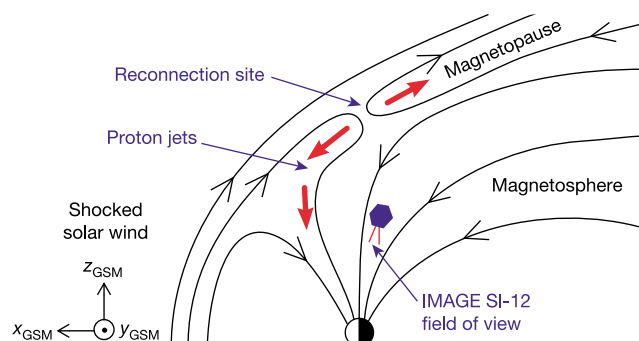


Figure 1 Diagram showing the link between reconnection at the magnetopause and its footprint in the ionosphere. The dayside proton auroral spot imaged by the IMAGE spacecraft (see Fig. 2) is created by precipitating solar wind protons that have crossed the high-latitude magnetopause and were accelerated by the magnetic kink resulting from magnetic reconnection between the solar-wind's and Earth's magnetic fields. Reconnection in this case happens between a northward (up in the figure) along the z axis of the geocentric solar magnetospheric (GSM) coordinate system directed solar-wind magnetic field and opposite magnetospheric field lines at the high-latitude location.

Models of reconnection are typically steady state. They describe what happens once the process is initiated and allowed to proceed indefinitely. Transient reconnection models have also been developed, which describe the consequence of artificially varying the reconnection rate, even including times when the rate drops to zero⁵. A key unanswered question, both theoretically and observationally, is how long the process can maintain itself naturally once initiated⁶. In other words, is the process intrinsically intermittent or continuous? Intermittent reconnection turns on and off. Continuous reconnection operates at a variable rate but never ceases; if the fluctuation is a small fraction of the average then the reconnection is classed as 'quasi-steady'. Attempts to deduce the steadiness of reconnection on the basis of *in situ* magnetopause observations were inconclusive⁷. Because signatures of reconnection (for example, plasma jets) are localized in the thin magnetopause, they can be observed only for a short time (a few minutes) when a spacecraft intersects the layer, even if the reconnection is actually continuous. Much of the published literature from ground-based or *in situ* measurements in the cusp has reported intermittent reconnection^{8–12}.

It has been established recently that the dayside proton auroral spot^{13,14} is created directly by solar-wind protons leaking through, and accelerated at, the magnetopause by way of reconnection between solar-wind and magnetospheric magnetic fields⁴ (Fig. 1). The proton aurora spot is detected by the SI-12 imager on the IMAGE spacecraft¹⁵. It detects the Doppler-shifted Lyman- α photons corresponding to precipitating charge-exchanged protons with energies of a few keV. As this is not the average energy of a typical proton distribution in the shocked solar wind, some sort of energization process is required. Magnetic reconnection at the magnetopause readily provides this energization⁴. Thus, the SI-12 imager is well suited to provide global images of the aurora created from precipitating protons that have been energized by magnetic reconnection at an active reconnection site.

Here we report two notable examples of long-duration proton auroral spot observations that imply continuous and even quasi-

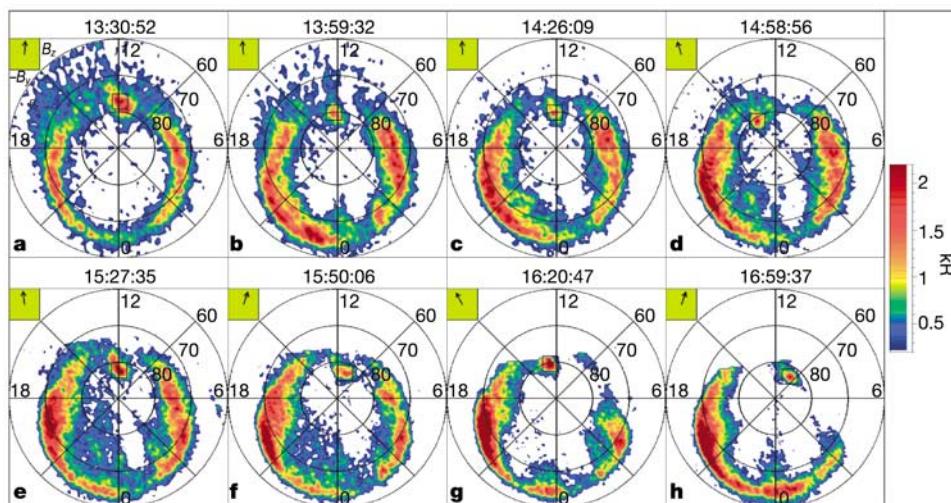


Figure 2 Snapshots of the proton aurora oval and spot on 18 March 2002 showing the continuous presence of the proton aurora spot. The Spectrographic Imager channel SI-12 on board the IMAGE spacecraft observes precipitating, high-energy protons¹⁵. On their way down along magnetic field lines, energetic protons temporarily capture electrons from the high-altitude atmospheric atoms and molecules. These newly formed hydrogen atoms may be in an excited state. When the atom decays to the ground state it releases a photon at the Lyman- α wavelength (121.567 nm). The SI-12 observes Doppler-shifted Lyman- α emission from precipitating protons with energy of several keV. The images are shown in a geomagnetic grid of latitudes and local time with the noon meridian at the top

and morning 06:00 h pointing to the right. The solar-wind magnetic field in $y_{\text{GSM}}-z_{\text{GSM}}$ plane is shown in the upper left inset, with north ($B_z > 0$) pointing up and east ($B_y > 0$) pointing to the left. The black square in each panel covers the $500 \times 500 \text{ km}^2$ area around the spot used in the computation of the mean brightness in Fig. 3c. The dayside proton aurora spot is seen uninterrupted over ~ 4 h. The spot appears on the dayside at $\sim 80^\circ$ latitude. Its location in magnetic local time (MLT) is correlated with the y -component of the solar-wind magnetic field, being in the pre-noon (post-noon) sector for negative (positive) B_y . Colour scale shows Lyman- α brightness in kilorayleighs, kR.

steady magnetopause reconnection for many hours. These observations were made during northward solar-wind magnetic field conditions, when the reconnection site is expected to be at the high-latitude magnetopause where the solar-wind and magnetospheric magnetic fields are antiparallel (Fig. 1).

The first reconnection event occurred on 18 March 2002 (Fig. 2). For the entire ~4-h interval shown, the solar-wind dynamic pressure (Fig. 3a) was constant (~17 nPa). The solar-wind magnetic field (B , Fig. 3b) was persistently northward directed ($B_z > 0$) while its y (east–west) component (B_y) fluctuated between positive and negative. Figures 2 and 3c–e show the uninterrupted presence of a bright proton auroral spot during this ~4-h interval. During a brief interval (5 min), the Cluster spacecraft by chance crossed the high-latitude magnetopause and observed proton jets accelerated by reconnection at the magnetopause⁴. The reconnection jets were observed on field lines that are linked to the spot in the ionosphere and with energy fluxes consistent with the spot brightness, thus providing direct evidence that the spot represents the remote signature of high-latitude magnetopause reconnection. Although Cluster could only observe reconnection for 5 min, the uninterrupted presence of the spot in the images in Fig. 2 implies continuous reconnection over ~4 h. Note that the peak and average brightness of the spot (Fig. 3c) remained high over the whole observation period, with some fluctuations.

After 14:40 UT, the magnetic local time (MLT) location of the spot (Figs 2, 3e) follows very closely the changes in the y component of the solar-wind magnetic field (Fig. 3b), with spot locations at pre-noon MLT for negative B_y and with post-noon locations with positive B_y (ref. 16). The correlation implies that reconnection is a directly driven process at the magnetopause and is consistent with antiparallel reconnection. That is, the reconnection site is located where the internal magnetospheric magnetic field is antiparallel to the external shocked solar-wind field^{14,17}. As the solar-wind magnetic field changes, the antiparallel reconnection site moves on the magnetopause. This is seen in the ionosphere as motion in MLT but not in latitude. This motion, combined with the near-constant brightness of the spot, indicates that the reconnection site is continuously active (not intermittent). Thus, viewed globally, the reconnection process never stops at the high-latitude magnetopause. Finally, as the spot moves in MLT in response to the solar wind B_y , it does not leave a long trail in latitude or in MLT, which would have indicated that proton aurorae are still created long (more than 4–5 min) after the cessation of reconnection.

The second continuous reconnection event occurred on 17/18 September 2000 (Fig. 4). This event does not have the added benefit of simultaneous *in situ* observations at the magnetopause. However, it is remarkable in that the dayside proton aurora spot was observed uninterruptedly for ~9 h (01:00–10:00 on 18 September), in addition to more than 3 h of continuous presence of the bright proton spot during the previous IMAGE apogee (northern hemisphere) pass (16:20–19:30 UT on 17 September). Before 05:45 UT on 18 September, similar to the previous event, the MLT location of the spot is well correlated with changes in the y -component of the solar-wind magnetic field, with similar implications in terms of antiparallel reconnection. Between 05:45 and 09:00 UT on 18 September, however, the solar-wind pressure and magnetic field direction were nearly constant. The proton spot brightness and location during the corresponding interval were also remarkably stable. Furthermore, the poleward boundary of the proton aurora spot (red line in Fig. 4d) was also rather stable. This stability rules out bursts of reconnection with periods longer than 4–5 min (the images are obtained on a two-minute cadence). Longer periods of reconnection bursts would produce observable brightness and size changes, slow periodic equatorward motions, and sudden poleward jumps of the proton spot poleward boundary.

The above observations can be summarized and further inter-

preted as follows. The dayside proton auroral spot, which recently has been shown to be the footprint of high-latitude reconnection, was observed continuously for close to 4 h in one event and more than 9 h in another. The continuous presence of the spot even as the solar-wind condition changed implies that continuous magnetic reconnection was occurring over these time spans. In a global sense, it is significant that reconnection at the high-latitude magnetopause never stops for any period of time longer than 4–5 min. For more stable solar-wind conditions in the second half of the second event, reconnection is continuous not only in a global sense, but its rate appears quasi-steady at the local magnetopause as well. The present finding explains the persistence of optical signatures of the cusp from ground-based observations during stable solar-wind magnetic field conditions^{18–20}.

For low-latitude reconnection during southward solar-wind magnetic field conditions, it was reported that the proton aurora could be seen on each reconnected field line for up to 5 min following reconnection²¹. In the present high-latitude reconnection case, this time might be longer. Any intermittency in the reconnection rate on 5 min or longer timescales should create observable signatures (changes in spot brightness and latitude) in the proton

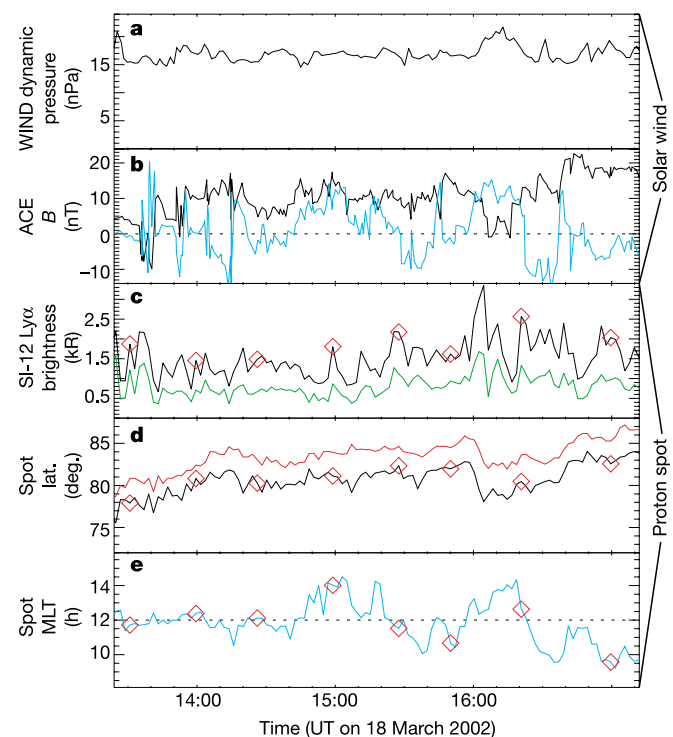


Figure 3 Continuous presence of dayside proton aurora spot (and implied reconnection) for ~4 h on 18 March 2002. **a**, The solar-wind dynamic pressure—measured by the WIND spacecraft and shifted in time by an average of 7 min to take into account the solar-wind convection time from the WIND position to the high-latitude magnetopause—was constant (~17 nPa) throughout this interval. **b**, The y (blue) and z (black) components of the solar-wind magnetic field measured by the ACE spacecraft, shifted in time by an average of 58 min. The field was mostly directed northward during the entire interval while the y -component is variable. **c**, IMAGE SI-12 detected peak (black) and mean (green) (averaged over $500 \times 500 \text{ km}^2$ area around the brightest pixel) Lyman- α brightness of the dayside proton auroral spot. Times of image snapshots shown in Fig. 2 are marked with red diamonds. **d**, **e**, The locations of the proton spot in magnetic latitude (black line in **d**) and MLT (blue line in **e**). In **d** we also show the location of the poleward boundary of the proton aurora spot (red line). Note the remarkable correlation between the spot MLT and the east–west (y) component of the solar-wind magnetic field (blue line in **b**) after 14:40 UT.

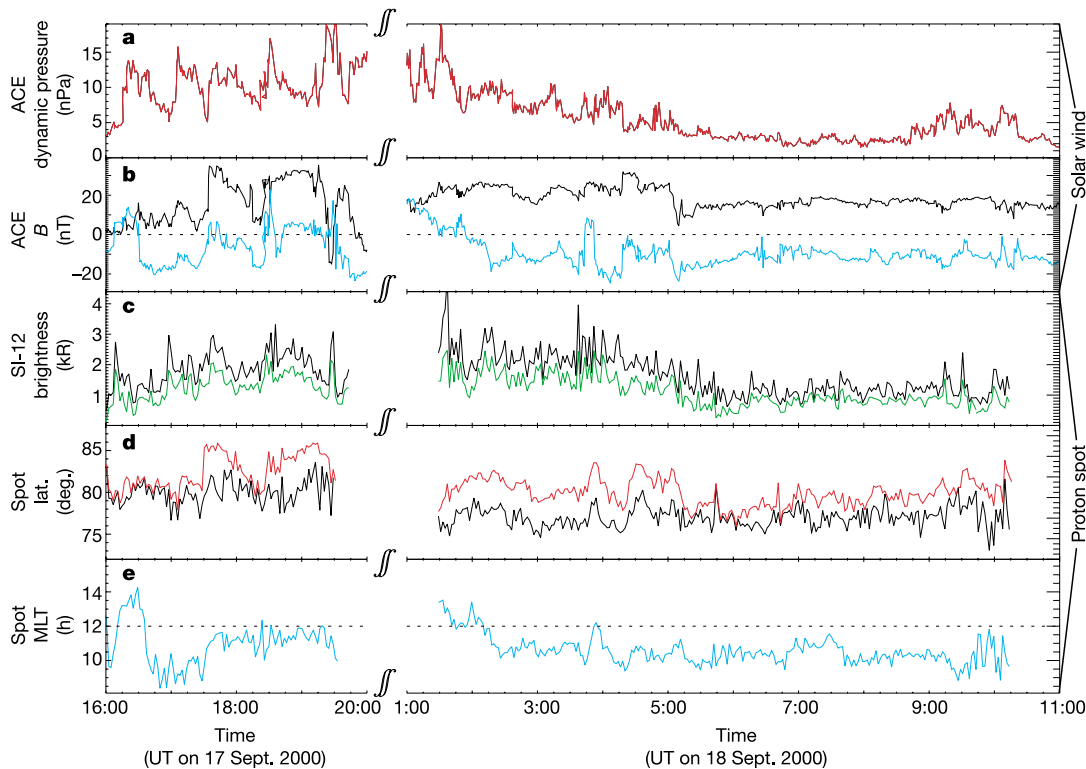


Figure 4 Continuous presence of dayside proton aurora spot (and implied reconnection) for ~9 h on 17/18 September 2000. The parameters are similar to Fig. 2. Note the remarkable steadiness of the spot brightness and location in magnetic latitude (**d**) and local time (**e**) from 05:45 to 09:00 UT on 18 September during steady solar-wind pressure (**a**) and magnetic field (**b**) conditions. At other times, the spot location is correlated with the

east–west (*y*) component of the solar-wind magnetic field, and its brightness is correlated with the solar-wind pressure. Between 19:30 and 01:20 UT there were no proton aurora observations by IMAGE because the spacecraft moved through perigee in the southern hemisphere.

aurora image sequences. Such changes are not seen in our 2-min-resolution images. Thus our findings are in contrast to reports favouring intermittent reconnection, which were made mostly under southward solar-wind magnetic field conditions^{8–10,12}. This contrast is more surprising considering that reconnection is thought to be more intermittent during northward solar-wind magnetic field, when the reconnection site is located at the high-latitude magnetopause (as opposed to a low-latitude site for southward solar-wind magnetic field) owing to the presence of fast shocked solar-wind flows²². A possible explanation for this discrepancy is the fact that most of the previous studies of reconnection have been made locally using ground-based or *in situ* measurements, which could not distinguish temporal variations of reconnection from motion of the reconnection site if the solar-wind magnetic field is not steady. At the local magnetopause, reconnection could well be intermittent if the solar-wind field changes.

The present results suggest that reconnection for northward solar-wind magnetic field is a directly driven process that is continuous when viewed in a global sense, or even in a local sense when the solar-wind conditions are steady. This regime of reconnection is in stark contrast to the storage and release of magnetic energy in explosive reconnection at the Sun (associated with solar flares) or in the magnetotail (associated with magnetospheric substorms). This contrast may suggest that, if driven persistently—in this case the shocked solar wind impinges on the magnetopause continuously—the process can be continuous. To evaluate fully the controlling factors that dictate the temporal behaviour of reconnection is a major objective of the planned NASA Magnetospheric Multiscale mission⁶.

Received 13 June; accepted 23 September 2003; doi:10.1038/nature02084.

- Cowley, S. W. H. in *Magnetic Reconnection in Space and Laboratory Plasmas* (ed. Hones, E. W.) 375–378 (Geophysics Monograph 30, American Geophysical Union, Washington DC, 1984).
- Hudson, H. & Ryan, J. High-energy particles in solar flares. *Annu. Rev. Astron. Astrophys.* **33**, 239–282 (1995).
- Angelopoulos, V. *et al.* Statistical characteristics of bursty bulk flow events. *J. Geophys. Res.* **103**, 21257–21280 (1994).
- Phan, T. *et al.* Simultaneous Cluster and IMAGE observations of cusp reconnection and auroral proton spot for northward IMF. *Geophys. Res. Lett.* **30**, doi:10.1029/2003GL016885 (2003).
- Scholer, M. Magnetic flux transfer at the magnetopause based on single X line bursty reconnection. *Geophys. Res. Lett.* **15**, 291–294 (1988).
- Report of the NASA Science and Technology Definition Team for the Magnetospheric Multiscale (MMS) Mission (NASA/TM-2000-209883, Goddard Space Flight Center, Greenbelt, MD, 1999).
- Phan, T. D. *et al.* Extended magnetic reconnection at the Earth's magnetopause from detection of bi-directional jets. *Nature* **404**, 848–850 (2000).
- Farrugia, C. J., Sandholt, P. E., Denig, W. F. & Torbert, R. B. Observation of a correspondence between poleward moving auroral forms and stepped cusp ion precipitation. *J. Geophys. Res.* **103**, 9309–9315 (1998).
- Lockwood, M. *et al.* Cusp ion steps, field-aligned currents and poleward moving auroral forms. *J. Geophys. Res.* **106**, 29555–29569 (2001).
- Lockwood, M., Davis, C. J., Onsager, T. G. & Scudder, J. D. Modelling signatures of pulsed magnetopause reconnection in cusp ion dispersion signatures seen at middle altitudes. *Geophys. Res. Lett.* **25**, 591–594 (1998).
- Sandholt, P. E. *et al.* Dynamic cusp aurora and associated pulsed reverse convection during northward interplanetary magnetic field. *J. Geophys. Res.* **105**, 12869–12894 (2000).
- Milan, S. E., Lester, M., Greenwald, R. A. & Sofko, G. The ionospheric signature of transient dayside reconnection and the associated pulsed convection return flow. *Ann. Geophys.* **17**, 1166–1171 (1999).
- Frey, H. U. *et al.* Proton aurora in the cusp. *J. Geophys. Res.* **A 107**, doi:10.1029/2001JA900161 (2002).
- Fuselier, S. A. *et al.* Cusp aurora dependence on IMF B_z. *J. Geophys. Res.* **A 107**, doi:10.1029/2001JA900165 (2002).
- Mende, S. B. *et al.* Far ultraviolet imaging from the IMAGE spacecraft: 3. Spectral imaging of Lyman alpha and OI 135.6 nm. *Space Sci. Rev.* **91**, 287–318 (2000).
- Newell, P. T., Meng, C.-I., Sibeck, D. G. & Lepping, R. Some low-altitude cusp dependencies on the interplanetary magnetic field. *J. Geophys. Res.* **94**, 8921–8927 (1989).
- Frey, H. U., Mende, S. B., Fuselier, S. A., Immel, T. J. & Ostgaard, N. Proton aurora in the cusp during southward IMF. *J. Geophys. Res.* **A 108**, doi:10.1029/2003JA009861 (2003).
- Mende, S. B., Bairden, R. L., Lanzarotti, L. J. & MacLennan, C. G. Magnetic impulses and associated optical signatures in the dayside aurora. *Geophys. Res. Lett.* **17**, 131–134 (1990).

19. Sandholt, P. E. IMF control of polar cusp and cleft auroras. *Adv. Space Res.* **8**, 21–34 (1988).
20. Newell, P. T. Do the dayside cusps blink? *Rev. Geophys. Suppl.* **33**, 665–668 (1995).
21. Lockwood, M. *et al.* IMF control of cusp proton emission intensity and dayside convection: Implications for component and anti-parallel reconnection. *Ann. Geophys.* **21**, 955–982 (2003).
22. Cowley, S. W. H. & Owen, C. J. A simple illustrative model of open flux tube motion over the dayside magnetopause. *Planet. Space Sci.* **37**, 1461–1475 (1989).

Acknowledgements We are indebted to the IMAGE team and J. L. Burch for the design and successful operation of the IMAGE mission.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to H.U.F. (hfrey@ssl.berkeley.edu).

Emergence of a molecular Bose–Einstein condensate from a Fermi gas

Markus Greiner¹, Cindy A. Regal¹ & Deborah S. Jin²

¹JILA, National Institute of Standards and Technology and Department of Physics, University of Colorado, ²Quantum Physics Division, National Institute of Standards and Technology, Boulder, Colorado 80309-0440, USA

The realization of superfluidity in a dilute gas of fermionic atoms, analogous to superconductivity in metals, represents a long-standing goal of ultracold gas research. In such a fermionic superfluid, it should be possible to adjust the interaction strength and tune the system continuously between two limits: a Bardeen–Cooper–Schrieffer (BCS)-type superfluid (involving correlated atom pairs in momentum space) and a Bose–Einstein condensate (BEC), in which spatially local pairs of atoms are bound together. This crossover between BCS-type superfluidity and the BEC limit has long been of theoretical interest, motivated in part by the discovery of high-temperature superconductors^{1–10}. In atomic Fermi gas experiments superfluidity has not yet been demonstrated; however, long-lived molecules consisting of locally paired fermions have been reversibly created^{11–15}. Here we report the direct observation of a molecular Bose–Einstein condensate created solely by adjusting the interaction strength in an ultracold Fermi gas of atoms. This state of matter represents one extreme of the predicted BCS–BEC continuum.

The basic idea behind this experiment is to start with a Fermi gas that has been evaporatively cooled to a high degree of quantum degeneracy, and adiabatically create molecules with a magnetic-field sweep across a Feshbach resonance. If the molecule creation conserves entropy and the initial atom gas is at sufficiently low temperature T compared to the Fermi temperature T_F , then the result should be a molecular sample with a significant condensate fraction^{13,16}. With a relatively slow sweep of an applied magnetic field that converts most of the fermionic atoms into bosonic molecules and an initial atomic gas below $T/T_F = 0.17$, we observe a molecular condensate in time-of-flight absorption images taken immediately following the magnetic-field sweep. The molecular condensate is not formed by any active cooling of the molecules, but rather merely by traversing the predicted BCS–BEC crossover regime.

Our experimental set-up and the procedure used to cool a gas of fermionic ⁴⁰K atoms to quantum degenerate temperatures have been detailed in previous work^{17,18}. In brief, after laser cooling and trapping we evaporatively cool the atoms in a magnetic trap. In

order to realize s-wave collisions in the ultracold Fermi gas, we use a mixture of atoms in two different spin states. For the final stage of evaporative cooling, the atoms are loaded into an optical dipole trap formed by a single far-red-detuned laser beam. The laser wavelength is $\lambda = 1,064$ nm, and the beam is focused to a waist of $15.5\ \mu\text{m}$. By lowering the depth of the optical trap, we evaporate the atomic gas to temperatures far below the Fermi temperature $T_F = (6N\nu_r^2\nu_z)^{1/3}h/k_B$. Here N is the particle number in each spin state, ν_r and ν_z are the radial and axial trap frequencies, h is Planck's constant, and k_B is Boltzmann's constant. For final radial trap frequencies between $\nu_r = 430$ Hz and 250 Hz and a fixed trap aspect ratio $\nu_r/\nu_z = 79 \pm 15$, we achieve temperatures between $0.36T_F$ and $0.04T_F$. (All temperatures of the Fermi gas given in this work are determined through surface fits to time-of-flight absorption images¹⁸.)

For this work we use a Feshbach resonance, which occurs when the energy of a quasibound molecular state becomes equal to the energy of two free atoms. The magnetic-field dependence of the resonance allows precise tuning of the atom–atom interaction strength in an ultracold gas¹⁹. Moreover, time-dependent magnetic fields can be used to reversibly convert atom pairs into extremely weakly bound molecules^{11–14,20–24}. The particular resonance used here is located at a magnetic field $B_0 = 202.1 \pm 0.1$ G and has a width of $w = 7.8 \pm 0.6$ G (refs 15, 25). The resonance affects collisions between atoms in the two lowest-energy spin states $|f = 9/2, m_f = -7/2\rangle$ and $|f = 9/2, m_f = -9/2\rangle$ of ⁴⁰K, where f denotes the total atomic angular momentum and m_f the magnetic quantum number.

To create bosonic molecules from the fermionic atoms, we first prepare an equal mixture of atoms in the $m_f = -9/2$ and $m_f = -7/2$ spin states at temperatures below quantum degeneracy. Then we apply a time-dependent sweep of the magnetic field starting above the Feshbach resonance value, where the atom interactions are effectively attractive, and ending below the resonance, where the atom interactions are effectively repulsive. In contrast to our previous work¹¹, the magnetic-field sweep is not only adiabatic with respect to the molecule creation rate, but also slow with respect to the collision rate and the radial trap frequency¹³. The magnetic field is typically ramped in 7 ms from $B = 202.78$ G to either $B = 201.54$ G or $B = 201.67$ G. With this magnetic-field sweep across the Feshbach resonance we convert between 78% and 88% of the atoms into molecules. To a very good approximation, these molecules have twice the polarizability of the atoms²⁶ and therefore are confined in the optical dipole trap with the same trapping frequency and twice the trap depth of the atoms. The molecules, which are all in the same internal quantum state, are highly vibrationally excited, very large in spatial extent, and extremely weakly bound. For a magnetic field 0.43 G below the Feshbach resonance ($B = 201.67$ G) the binding energy $\hbar^2/ma^2$ is 8 kHz, where m is the atomic mass and $2\pi\hbar$ is Planck's constant. The molecule size, which we estimate as $a/2$, is $\sim 1,650a_0$, where a_0 is the Bohr radius and a is the atom–atom scattering length given by $a = 174a_0[1 + w/(B_0 - B)]$ (ref. 18). At this magnetic field, the molecule size is one order of magnitude smaller than the calculated intermolecular distance.

A critical element of this experiment is that the lifetime of these weakly bound molecules can be much longer than the typical collision time in the gas and longer than the radial trapping period^{12–15}. In previous work, we found that the ⁴⁰K₂ molecule lifetime increases dramatically near the Feshbach resonance and reaches ~ 100 ms at a magnetic field 0.43 G below the Feshbach resonance for a peak density of $n_{\text{pk}} = 1.5 \times 10^{13}\ \text{cm}^{-3}$ (ref. 15). It is predicted that this increased molecule lifetime only occurs for dimers of fermionic atoms²⁷. The relatively long molecule lifetime near the Feshbach resonance allows the atom/molecule mixture to achieve thermal equilibrium during the magnetic-field sweep. Note, however, that the large aspect ratio of the optical trap gives rise to a

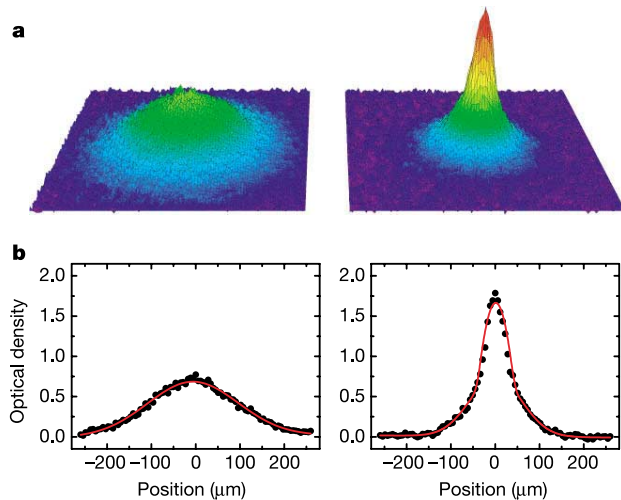


Figure 1 Time-of-flight images of the molecular cloud, taken with a probe beam along the axial direction after 20 ms of free expansion. Data are shown for temperatures above and below the critical temperature for Bose–Einstein condensation. **a**, Surface plot of the optical density for a molecule sample created by applying a magnetic-field sweep to an atomic Fermi gas with an initial temperature of $0.19T_F$ ($0.06T_F$) for the left (right) image. Here the radial trapping frequency of the optical trap was 350 Hz (260 Hz). When we start with the lower initial temperature of the fermionic atoms (right) and ramp across the Feshbach resonance from $B = 202.78$ G to 201.54 G in 10 ms, the molecules form a BEC. During expansion the interparticle interaction was reduced by rapidly moving the magnetic field 4 G further away from the Feshbach resonance. The total molecule number was 470,000 (200,000) for the left (right) picture. The surface plots are the averages of ten images. **b**, Cross-sections through images corresponding to the parameters given above (dots), along with bimodal surface fits (lines). The fits yield no condensate fraction and a temperature of $T = 250$ nK $= 0.90T_c$ for the left graph, and a 12% condensate fraction and a temperature of the thermal component of $T = 79$ nK $= 0.49T_c$ for the right graph. Here, T_c is the calculated critical temperature for a non-interacting BEC in thermal equilibrium.

strongly anisotropic system. Thus for the relatively short timescale of the experiments reported here we may attain only local equilibrium in the axial direction²⁸.

To study the resulting atom–molecule mixture after the magnetic-field sweep, we measure the momentum distribution of both the molecules and the residual atoms using time-of-flight absorption imaging. After typically 10–20 ms of expansion, we apply a radio frequency (r.f.) pulse that dissociates the molecules into free atoms in the $m_f = -5/2$ and $m_f = -9/2$ spin states¹¹. Immediately after this r.f. dissociation pulse, we take a spin-selective absorption image. The r.f. pulse has a duration of 140 μs and is detuned 50 kHz beyond the molecule dissociation threshold so that it does not affect the residual unpaired atoms in the $m_f = -7/2$ state. We selectively detect the expanded molecule cloud by imaging atoms transferred by the r.f. dissociation pulse into the previously unoccupied $m_f = -5/2$ state. Alternatively, we can image only the expanded atom cloud by detecting atoms in the $m_f = -7/2$ spin state.

Close to the Feshbach resonance, the atoms and molecules are strongly interacting with effectively repulsive interactions. As shown by Petrov *et al.* (ref. 27 and references therein), the scattering length for atom–molecule and molecule–molecule collisions close to the Feshbach resonance are $1.2a$ and $0.6a$ respectively. During the initial stage of expansion, the positive interaction energy is converted into additional kinetic energy of the expanding cloud. Therefore the measured momentum distribution is very different from the original momentum distribution of the trapped cloud. In order to reduce the effect of these interactions on the molecule time-of-flight images, we use the magnetic-field Feshbach resonance to

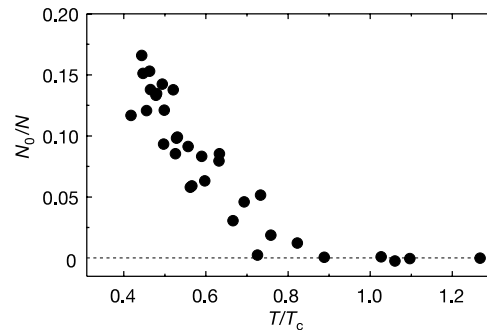


Figure 2 Molecular condensate fraction N_0/N versus the scaled temperature T/T_c . The temperature of the molecules is varied by changing the initial temperature of the fermionic atoms before the formation of the molecules. All other parameters are similar to those described in Fig. 1 legend. We observe the onset for Bose–Einstein condensation at a temperature of $\sim 0.8T_c$; the dashed line marks zero condensate fraction.

control the interparticle interaction strength during expansion. We can significantly reduce the interaction energy momentum kick by rapidly changing the magnetic field before we switch off the optical trap for expansion. The field is lowered typically by 4 G in 10 μs. At this magnetic field further away from the resonance, a is reduced to $\sim 500a_0$. We find that this magnetic-field jump results in a loss of typically 50% of the molecules, which we attribute to the reduced molecule lifetime away from the Feshbach resonance.

Below an initial temperature of $0.17T_F$, we observe the sudden onset of a pronounced bimodal momentum distribution for the molecules. Figure 1 shows such a bimodal distribution for an experiment starting with an initial temperature of $0.06T_F$; for comparison, we also show the resulting molecule momentum distribution for an experiment starting at $0.19T_F$. The bimodal momentum distribution is a striking indication that the cloud of weakly bound molecules has undergone a phase transition to a BEC^{29–31}.

To obtain thermodynamic information about the molecule cloud, we fitted the momentum distribution with a two-component fit. The fit function is the sum of an inverted parabola (describing the Thomas–Fermi momentum distribution of a bosonic condensate) and a gaussian momentum distribution (describing the non-condensed component of the molecule cloud). In Fig. 2 the measured condensate fraction is plotted as a function of the fitted temperature of the thermal component in units of the critical temperature for an ideal Bose gas, $T_c = 0.94(Nv_F^2v_z)^{1/3}h/k_B$. Here N is the total number of molecules when there is no change of the magnetic field for the expansion. Note that this measurement may underestimate the original condensate fraction owing to loss of molecules during expansion. From Fig. 2 we determine an actual critical temperature for the interacting molecules and for our trap geometry of $0.8 \pm 0.1T_c$. Such a decrease of the critical temperature relative to the ideal gas prediction is expected owing to repulsive interactions in a trapped gas³².

We find that the creation of a BEC of molecules requires that the Feshbach resonance be traversed sufficiently slowly. This is illustrated in Fig. 3, where the measured condensate fraction is plotted versus the ramp time across the Feshbach resonance, starting with a Fermi gas at a temperature $0.06T_F$. Our fastest sweeps result in a much smaller condensate fraction, whereas the largest condensate fraction appears for a B -field sweep of 3–10 ms. For even slower magnetic-field sweeps, we find that the condensate fraction slowly decreases. We attribute this effect to a finite lifetime of the condensate. Note that the timescale of the experiment is short compared to the axial trap frequency. Therefore the condensate may not have global phase coherence in the axial direction²⁸. The inset of

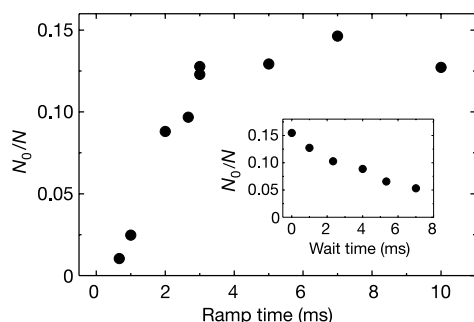


Figure 3 Dependence of condensate formation on magnetic-field sweep rate and measurement of condensate lifetime. We plot the fraction of condensed molecules versus the time in which the magnetic field is ramped across the Feshbach resonance from 202.78 G to 201.54 G. The condensate fraction is measured after an additional waiting time of 1 ms. The initial atom gas temperature is $0.06T_F$, the total molecule number is 150,000, and the final radial trap frequency is 260 Hz. For the full range of ramp times the number of molecules created remains constant. Inset, plot of condensate fraction versus the wait time after a 10-ms magnetic-field ramp. The molecule number is not significantly reduced on this timescale, and the lifetime of the condensate is instead determined by a heating rate, which we measure to be 3 ± 1 nK ms $^{-1}$. This heating rate is presumably due to density-dependent inelastic loss processes.

Fig. 3 shows a plot of the lifetime of the condensate. The observed reduction in condensate fraction is accompanied by heating of the molecule gas, presumably due to the density-dependent collisional decay of molecules into more tightly bound states.

Rapidly changing the interaction strength for time-of-flight expansion of the condensate allows us to measure the interaction energy in the molecular sample. Figure 4 shows a plot of the expansion energy of the molecule BEC for various interaction strengths during time-of-flight expansion. Here the condensate is created at a fixed interaction strength, and thus the initial peak density n_{pk} is constant. The data show that the expansion energy is proportional to a . The linear dependence suggests that the molecule–molecule scattering length is proportional to the atom–atom scattering length, as predicted in ref. 27. In addition, we find that the expansion energy extrapolates to near-zero energy for $a = 0$. This is consistent with a trapped BEC with relatively strong repulsive interactions. Assuming the molecule–molecule interaction strength calculated in ref. 27, this measurement allows us to determine the peak density of the interacting condensate as $n_{pk} = 7 \times 10^{12}$ cm $^{-3}$.

A fundamental aspect of our experiment is that we start with a quantum degenerate Fermi gas of atoms. The BEC, which is observed immediately after the magnetic field is ramped across the Feshbach resonance, therefore requires a drastic change of the quantum statistical thermodynamics of the gas. This change is not due to evaporative cooling, and the total number of atoms (adding both free atoms and those bound in molecules) is conserved by the field sweep. In Fig. 5 we show the dependence of the condensate fraction on the initial temperature of the Fermi gas. We find that a BEC is formed when the initial temperature is below $0.17T_F$. If we assume that entropy is conserved in the sweep across the Feshbach resonance, then creating the molecular BEC depends on starting with a Fermi gas at sufficiently low T/T_F to give low initial entropy¹⁶. At the onset of Bose–Einstein condensation, our temperature measurement indicates a modest 40% increase in the total entropy after the magnetic-field sweep, estimated from an ideal gas model. The inset in Fig. 5 compares the absolute temperature of atoms and molecules after the magnetic-field sweep. For the molecules, the temperature is determined by a fit to the non-condensate fraction. We find that atoms and molecules are well thermalized. Note that the atoms and molecules are not in full chemical equilibrium¹³. Even

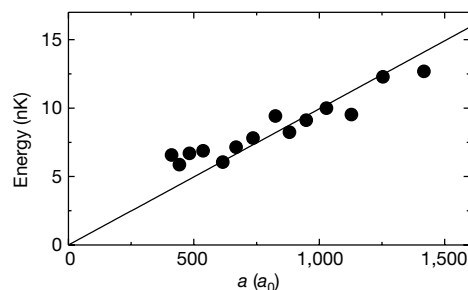


Figure 4 Total expansion energy per particle for the molecular condensate versus the interaction strength during expansion. As the molecular condensate is released from the trap ($\nu_r = 260$ Hz), collisions convert the mean-field interaction energy into kinetic energy of the expanding molecules. In this measurement, the BEC is created in a regime for which we calculate the atom–atom scattering length to be $3,300a_0$. For expansion, the magnetic field is rapidly changed to different final values. The graph shows the expansion energy of the condensate fraction determined from a bimodal fit versus the atom–atom scattering length a corresponding to the magnetic field during expansion. The total molecule number is 140,000, the magnetic field before expansion is 201.54 G, and we measure the condensate fraction to be 14%. The line is a linear fit with no offset. We find that the kinetic energy of the condensate molecules is proportional to a .

though the final binding energy of the molecules is significantly larger than $k_B T$, we only observe conversion efficiencies of up to 88%. To study the reversibility of slow ramps (10-ms ramp time) across the Feshbach resonance, we have ramped the magnetic field back to the attractive side of the resonance after creating a molecular condensate and then measured the temperature of the resulting Fermi gas. We find that the gas is heated by 27 ± 7 nK in this double ramp, independent of the initial temperature.

In conclusion, we have created a BEC of weakly bound molecules starting with a gas of ultracold fermionic atoms. The molecular BEC has been detected through a bimodal momentum distribution, and effects of the strong interparticle interaction have been investigated. The molecular BEC reported here, which appears on the repulsive side of the Feshbach resonance, is related in a continuous way to BCS-type fermionic superfluidity on the attractive side of the resonance. Our experiment corresponds to the BEC limit, in which superfluidity occurs owing to Bose–Einstein condensation of essentially local pairs whose binding energy is much larger than the Fermi energy. The dimensionless parameter $1/k_F a$, which drives the crossover from a BCS-superfluid to a molecular BEC^{4,5}, is thus

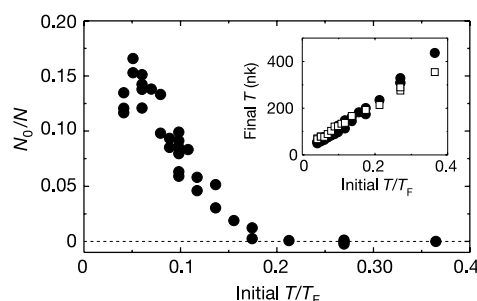


Figure 5 Dependence of the condensate fraction and the temperature of the atom–molecule mixture on the initial scaled temperature T/T_F of the Fermi gas. The condensate fraction is plotted versus T/T_F of the fermionic atoms before the magnetic sweep. In the inset, the temperature of the atoms (open boxes) and the thermal fraction of the molecules (closed circles) are plotted versus T/T_F before the sweep. This is the same data set as in Fig. 2; the dashed line marks zero condensate fraction.

positive and large compared to one. In contrast, near the Feshbach resonance where $1/k_F a$ goes through zero, the system may be described neither by a BEC of molecular dimers nor by a BCS-state of correlated pairs in momentum space, a situation which has been termed ‘resonance superfluidity’⁶. Indeed, our experiment passes through this unexplored crossover regime, and with initial temperatures below $0.1 T_F$ the Fermi gas is well below the predicted critical temperature in the crossover regime of $0.52 T_F$ (ref. 10). In future work, it will be interesting to investigate this system on the attractive side of the resonance and look for evidence of fermionic superfluidity. □

Received 3 November; accepted 14 November 2003; doi:10.1038/nature02199.
Published online 26 November 2003.

- Leggett, A. J. Cooper pairing in spin-polarized Fermi systems. *J. Phys. C (Paris)* **41**, 7–19 (1980).
- Noziers, P. & Schmitt-Rink, S. Bose condensation in an attractive fermion gas: From weak to strong coupling superconductivity. *J. Low-Temp. Phys.* **59**, 195–211 (1985).
- Drechsler, M. & Zwerger, W. Crossover from BCS-superconductivity to Bose-condensation. *Ann. Phys.* **1**, 15–23 (1992).
- Haussmann, R. Properties of a Fermi-liquid at the superfluid transition in the crossover region between BCS superconductivity and Bose-Einstein condensation. *Phys. Rev. B* **49**, 12975–12983 (1994).
- Randeria, M. in *Bose-Einstein Condensation* (eds Griffin, A., Snoke, D. W. & Stringari, S.) 355–392 (Cambridge Univ. Press, Cambridge, UK, 1995).
- Holland, M., Kokkelmans, S., Chiofalo, M. L. & Walser, R. Resonance superfluidity in a quantum degenerate Fermi gas. *Phys. Rev. Lett.* **87**, 120406 (2001).
- Timmermans, E., Furuya, K., Milonni, P. W. & Kerman, A. K. Prospect of creating a composite Fermi-Bose superfluid. *Phys. Lett.* **285**, 228–233 (2001).
- Ohashi, Y. & Griffin, A. BCS-BEC crossover in a gas of Fermi atoms with a Feshbach resonance. *Phys. Rev. Lett.* **89**, 130402 (2002).
- Milstein, J. N., Kokkelmans, S. & Holland, M. J. Resonance theory of the crossover from Bardeen-Cooper-Schrieffer superfluidity to Bose-Einstein condensation in a dilute Fermi gas. *Phys. Rev. A* **66**, 043604 (2002).
- Ohashi, Y. & Griffin, A. Superfluid transition temperature in a trapped gas of Fermi atoms with a Feshbach resonance. *Phys. Rev. A* **67**, 033603 (2003).
- Regal, C. A., Ticknor, C., Bohn, J. L. & Jin, D. S. Creation of ultracold molecules from a Fermi gas of atoms. *Nature* **424**, 47–50 (2003).
- Strecker, K. E., Partridge, G. B. & Hulet, R. G. Conversion of an atomic Fermi gas to a long-lived molecular Bose gas. *Phys. Rev. Lett.* **91** (2003).
- Cubizolles, J., Bourdel, T., Kokkelmans, S. J. J. M. F., Shlyapnikov, G. V., Salomon, C. Production of long-lived ultracold Li_2 molecules from a Fermi gas. Preprint at (<http://arXiv.org/cond-mat/0308018>) (2003).
- Jochim, S. *et al.* Pure gas of optically trapped molecules created from fermionic atoms. Preprint at (<http://arXiv.org/cond-mat/0308095>) (2003).
- Regal, C. A., Greiner, M., & Jin, D. S. Lifetime of molecule-atom mixtures near a Feshbach resonance in ^{40}K . Preprint at (<http://arXiv.org/cond-mat/0308606>) (2003).
- Carr, L. D., Shlyapnikov, G. V. & Castin, Y. Achieving a BCS transition in an atomic Fermi gas. Preprint at (<http://arXiv.org/cond-mat/0308306>) (2003).
- DeMarco, B. & Jin, D. S. Onset of Fermi degeneracy in a trapped atomic gas. *Science* **285**, 1703–1706 (1999).
- Regal, C. A. & Jin, D. S. Measurement of positive and negative scattering lengths in a Fermi gas of atoms. *Phys. Rev. Lett.* **90**, 230404 (2003).
- Stwalley, W. C. Stability of spin-aligned hydrogen at low temperatures and high magnetic fields: New field-dependent scattering resonances and predissociations. *Phys. Rev. Lett.* **37**, 1628–1631 (1976).
- Donley, E. A., Claussen, N. R., Thompson, S. T. & Wieman, C. E. Atom-molecule coherence in a Bose-Einstein condensate. *Nature* **417**, 529–533 (2002).
- Chin, C., Kerman, A. J., Vuletic, V. & Chu, S. Sensitive detection of cold cesium molecules formed on Feshbach resonances. *Phys. Rev. Lett.* **93**, 033201 (2003).
- Herbig, J. *et al.* Preparation of a pure molecular quantum gas. *Science* **301**, 1510–1513 (2003).
- Durr, S., Volz, T., Marte, A. & Rempe, G. Observation of molecules produced from a Bose-Einstein condensate. Preprint at (<http://arXiv.org/cond-mat/0307440>) (2003).
- Xu, K. *et al.* Formation of quantum-degenerate sodium molecules. Preprint at (<http://arXiv.org/cond-mat/0310027>) (2003).
- Loftus, T., Regal, C. A., Ticknor, C., Bohn, J. L. & Jin, D. S. Resonant control of elastic collisions in an optically trapped Fermi gas of atoms. *Phys. Rev. Lett.* **88**, 173201 (2002).
- Ratcliff, L. B., Fish, J. L. & Konowalow, D. D. Electronic transition dipole moment functions for transitions among the twenty-six lowest-lying states of Li_2 . *J. Mol. Spectrosc.* **122**, 293–312 (1987).
- Petrov, D. S., Salomon, C. & Shlyapnikov, G. V. Weakly bound dimers of fermionic atoms. Preprint at (<http://arXiv.org/cond-mat/0309010>) (2003).
- Shvachuck, I. *et al.* Bose-Einstein condensation into non-equilibrium states studied by condensate focusing. *Phys. Rev. Lett.* **89**, 270–404 (2002).
- Anderson, M. H., Ensher, J. R., Matthews, M. R., Wieman, C. E. & Cornell, E. A. Observation of Bose-Einstein condensation in a dilute atomic vapor. *Science* **269**, 198–201 (1995).
- Davis, K. B. *et al.* Bose-Einstein condensation in a gas of sodium atoms. *Phys. Rev. Lett.* **75**, 3969–3973 (1995).
- Bradley, C. C., Sackett, C. A. & Hulet, R. G. Bose-Einstein condensation of lithium: observation of limited condensate number. *Phys. Rev. Lett.* **78**, 985–989 (1997).
- Giorgini, S., Pitaevskii, L. P. & Stringari, S. Condensate fraction and critical temperature of a trapped interacting Bose gas. *Phys. Rev. A* **54**, R4633–R4636 (1996).

Acknowledgements We thank L. D. Carr, E. A. Cornell, C. E. Wieman, W. Zwerger and I. Bloch for discussions, and J. Smith for experimental assistance. This work was supported by NSF and NIST. C.A.R. acknowledges support from the Hertz Foundation.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to M.G. (markus.greiner@colorado.edu).

Direct observation of Tomonaga–Luttinger-liquid state in carbon nanotubes at low temperatures

Hiroyoshi Ishii¹, Hiromichi Kataura¹, Hidetsugu Shiozawa¹, Hideo Yoshioka², Hideo Otsubo¹, Yasuhiro Takayama¹, Tsuneaki Miyahara¹, Shinzo Suzuki¹, Yohji Achiba¹, Masashi Nakatake³, Takamasa Narimura¹, Mitsuharu Higashiguchi⁵, Kenya Shimada⁶, Hirofumi Namatame⁶ & Masaki Taniguchi^{4,6}

¹Graduate School of Science, Tokyo Metropolitan University, Minami-Ohsawa 1-1, Hachioji, Tokyo 192-0397, Japan

²Department of Physics, Nara Women's University, Nara 630-8506, Japan

³Photon Factory, High Energy Accelerator Research Organization, Tsukuba 305-0801, Japan

⁴Graduate School of Science, ⁵Department of Physical Science, ⁶Hiroshima Synchrotron Radiation Center, Hiroshima University, Higashi-Hiroshima, 739-8526, Japan

The electronic transport properties of conventional three-dimensional metals are successfully described by Fermi-liquid theory. But when the dimensionality of such a system is reduced to one, the Fermi-liquid state becomes unstable to Coulomb interactions, and the conduction electrons should instead behave according to Tomonaga–Luttinger-liquid (TLL) theory. Such a state reveals itself through interaction-dependent anomalous exponents in the correlation functions, density of states and momentum distribution of the electrons^{1–3}. Metallic single-walled carbon nanotubes (SWNTs) are considered to be ideal one-dimensional systems for realizing TLL states^{4–6}. Indeed, the results of transport measurements on metal–SWNT and SWNT–SWNT junctions have been attributed^{7–9} to the effects of tunnelling into or between TLLs, although there remains some ambiguity in these interpretations¹⁰. Direct observations of the electronic states in SWNTs are therefore needed to resolve these uncertainties. Here we report angle-integrated photoemission measurements of SWNTs. Our results reveal an oscillation in the π -electron density of states owing to one-dimensional van Hove singularities, confirming the one-dimensional nature of the valence band. The spectral function and intensities at the Fermi level both exhibit power-law behaviour (with almost identical exponents) in good agreement with theoretical predictions for the TLL state in SWNTs.

The single-particle spectral function $\rho(\omega)$ of TLL states is predicted^{1–3} to have a power-law dependence on the binding energy ω near the Fermi energy E_F ($\rho(\omega) \propto |\omega|^\alpha$). As photoemission spectroscopy can directly measure the spectral profile near E_F , photoemission studies on quasi-one-dimensional (1D) conductors and artificial quantum wires have been extensively performed^{2,11–17}. In fact, the suppression in photoemission intensity near E_F was observed; many photoemission data suggested that a TLL state may be realized in quasi-1D systems. However, the observed exponents α ($\alpha > 1$) were much larger than the theoretical upper

limit in the single-band Hubbard model, $\alpha = 1/8$ (refs 1–3). To understand this discrepancy, extrinsic or intrinsic origins, such as opening of pseudogaps produced by inelastic processes, inhomogeneity of the sample surface, formation of charge or spin density wave gaps, and the presence of three-dimensional (3D) coupling, have been discussed^{11–17}. In SWNTs, it has been shown theoretically that the instability due to a lattice distortion, which generally occurs in a 1D system, is not important^{6,18}, and that a SWNT can be regarded as an ideal 1D system along the direction of its axis^{4–6}. It is of interest to investigate directly the electronic states of SWNTs by photoemission spectroscopy. However, the interpretation of photoemission data is not straightforward, because SWNTs in fact form a bundle. In addition to the effects discussed in previous work, we must carefully consider the influence of the intertube interaction on spectral profiles.

In SWNTs, the electronic states along the circumference are quantized so that peak structures due to 1D van Hove singularities (VHS) appear in the density of states^{18,19}. However, using photoemission spectroscopy, no one to date has succeeded in observing even the VHS structures^{20,21}. In photoemission measurements, it is necessary to prepare a relatively large sample consisting of SWNTs with uniform diameters. We prepared SWNT samples by the laser vaporization method²² (see Methods for details). Figure 1 shows images of SWNTs taken by transmission electron microscopy (TEM). Bundles consisting of several tens of SWNTs are randomly

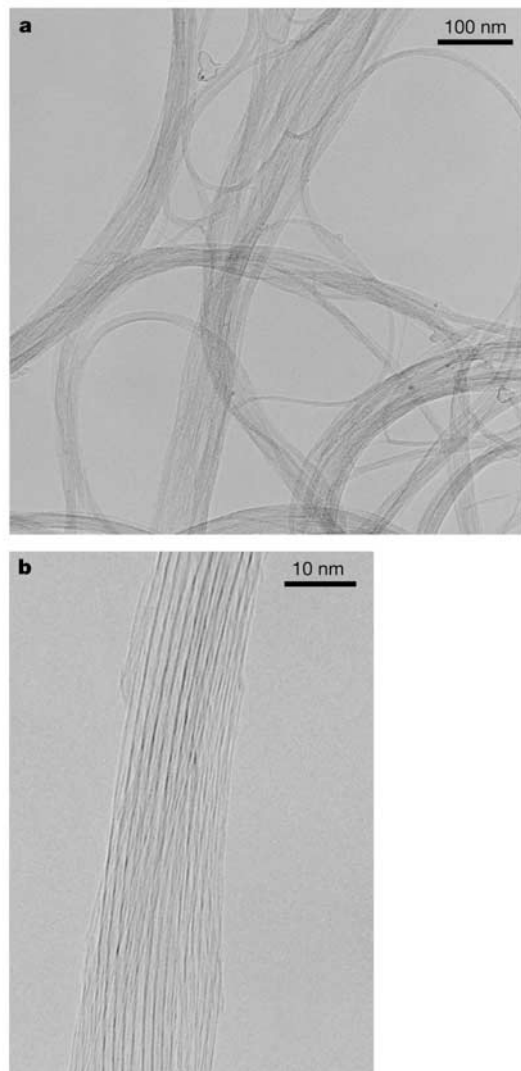


Figure 1 TEM images of SWNTs. **a**, Bulk sample; **b**, a SWNT bundle.

oriented. The photoemission experiments were performed using synchrotron radiation; the experimental set-ups are described in Methods.

Figure 2 shows the photoemission spectra of SWNT samples and graphite measured with an energy resolution of 50 meV. As shown in the Fig. 2 inset, the intense structure around the 8-eV peak and the structure around the 3-eV peak are due to σ and π bands, respectively. The broad trapezoid structure around 19 eV is the C 2s band. The shapes and positions of these structures are consistent with the results of refs 20 and 21. The valence band spectrum of the SWNT is similar as a whole to that of graphite. The difference in spectral features between the SWNT and graphite appears in the energy region near E_F . For SWNT samples, three peak structures (denoted by S_1 , S_2 and M_1 in Fig. 2) can be clearly seen around the binding energies of 0.45, 0.75 and 1.0 eV, respectively. On the other hand, the spectrum of graphite shows no structure in this energy region.

First, we compare the photoemission spectra with the optical absorption spectrum. In Fig. 2, the absorption spectrum was plotted with the energy reduced to 50%, because the optical transitions occur between mirror image spikes of density of states owing to the 1D VHS^{23,24}. Furthermore, the spectrum was shifted towards a higher binding energy by 0.1 eV, because an energy shift of the electronic structure due to doping of the SWNTs by the Cu substrate occurs, as reported in scanning tunnelling spectroscopy¹⁹. As can be seen from the figure, the S_2 and M_1 peaks in the photoemission spectrum correspond to the 1.3- and 1.8-eV peaks, respectively, in the optical spectrum. On the other hand, the S_1 peak was observed at a lower binding energy than the 0.7-eV peak in the optical spectrum. The discrepancy is consistent with the observation that the lowest optical transition around the fundamental absorption edge is shifted towards a higher photon energy²⁴.

We then compared the photoemission spectra with the calculated

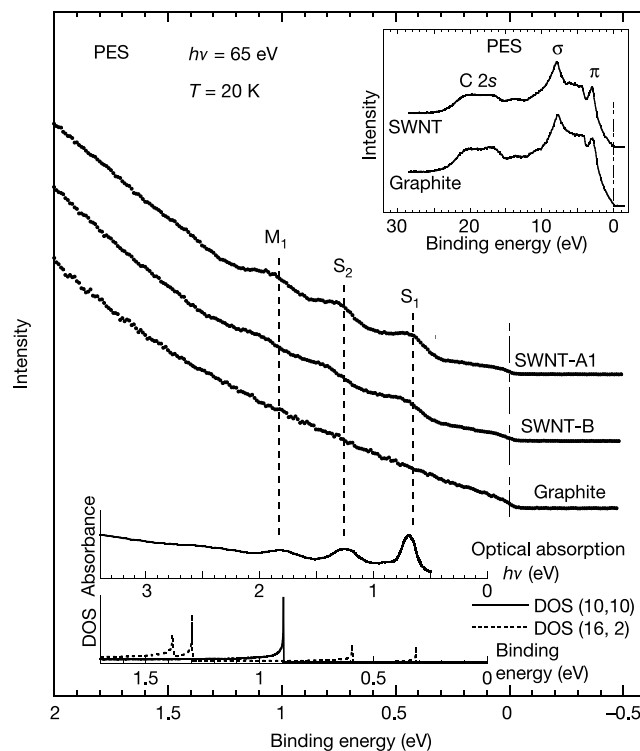


Figure 2 Photoemission spectra of SWNTs and graphite near E_F measured at $h\nu = 65$ eV. The optical absorption spectrum and the calculated densities of states (DOS) for SWNTs with chiral indices (10, 10) and (16, 2) are also shown. Inset, valence band photoemission spectra (PES), which were measured at KEK-PF.

π band structure based on the tight-binding model¹⁸. The structure and electronic properties of SWNTs can be defined by a pair of integers, the so-called chiral index expressed as (n_1, n_2) ; when $(2n_1 + n_2)/3$ is an integer, the SWNTs are metallic, and the others are semiconducting¹⁸. For the SWNT-A1 sample, the mean diameter of the SWNTs was estimated to be 1.37 nm from Raman spectra (Methods). For SWNTs with diameters near 1.37 nm, for example, a metallic SWNT with chiral index (10, 10) has a diameter of 1.375 nm, and a semiconducting SWNT with (16, 2) has a diameter of 1.357 nm. The densities of states for the SWNTs with these chiral indices (S. Maruyama, personal communication) are also shown in Fig. 2. As can be seen from the figures, the S_1 and S_2 peaks and the M_1 peak correspond to the structures of semiconducting and metallic SWNTs, respectively. From the above comparison, the three observed peak structures are due to the spikes caused by the 1D VHS in the occupied density of states.

As can be seen from Fig. 2, distinct peaks were not observed at binding energies above 1.2 eV, and the three observed peaks were broad in comparison with the spikes in the calculated π band of the SWNT. The diameter distribution of SWNTs in the sample makes the photoemission peaks unclear. Ichida *et al.* analysed the optical spectra of SWNTs taking into account their diameter distribution²⁴. Using their procedure, we calculated the photoemission spectrum of the SWNT-A1 sample. We assumed that the diameters of SWNTs were distributed according to a gaussian function with a standard deviation (Γ_d) around a mean value of 1.37 nm (Fig. 3 inset). The total density of states of the sample, $D(E)$, was obtained using the density of states for each SWNT calculated using the overlap integral $\gamma = 2.9$ eV (S. Maruyama, personal communication). The spectrum was calculated by smearing out $D(E)$ with a gaussian curve and with a lorentzian curve. The details of the fitting procedure are described in Methods. The line of best fit calculated using $\Gamma_d = 0.1$ nm, $\Gamma = 20$ meV (the spectral lifetime broadening) and

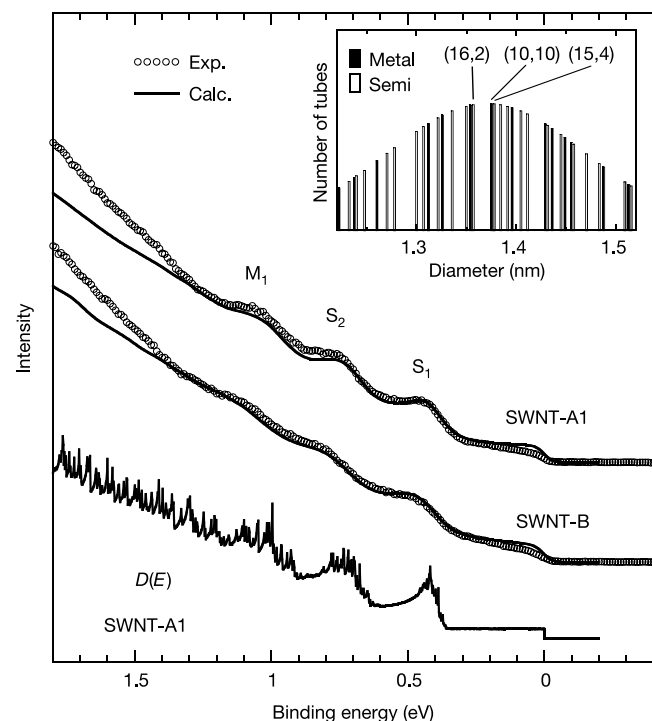


Figure 3 Comparison between the experimental and calculated spectra of the SWNT-A1 and SWNT-B samples. The density of states $D(E)$, calculated taking into account the diameter distribution of SWNTs, is indicated by a solid line. The inset shows the assumed diameter distribution for the SWNT-A1 sample. The sample consists of metallic SWNTs (filled bars) and semiconducting SWNTs (open bars).

$E_S = 0.1$ eV (the amount of the energy shift) is represented by a solid line in Fig. 3. The energy positions and shapes of the three peaks were well reproduced by the calculated spectrum. For the SWNT-B sample with a mean tube diameter of 1.25 nm (Fig. 2), the peak structures appeared at higher binding energies than the respective corresponding structures of the SWNT-A1 sample. The spectrum was well reproduced by the spectrum calculated using $\Gamma_d = 0.14$ nm, $\Gamma = 25$ meV and $E_S = 0.1$ eV (Fig. 3). The difference in peak position comes from the difference in diameter.

The observed spectrum is different in intensity from the calculated spectrum in binding energies above 1.3 eV (Fig. 3). This difference is due to the overlap with the tail of the σ band. Another difference can be seen near E_F . The observed photoemission intensity decreases rapidly near E_F compared with the calculated intensity; the deviation from the calculated spectrum increases with decreasing binding energy. To see the detailed spectral features near E_F , high-resolution photoemission spectra of the SWNT-A2 sample were measured with an energy resolution of 13 meV at $h\nu = 30$ eV (Fig. 4). In the spectrum measured at 10 K, the intensity near E_F is strongly suppressed; the tailing structure that appeared above E_F grows with increasing temperature. The intensity at E_F increases with increasing temperature. In the left inset of Fig. 4, we plot the spectra measured with an energy resolution of 15 meV at $h\nu = 65$ eV as a function of the binding energy on a double-logarithmic scale. The spectra do not show the photon energy dependence. Figure 5 shows the ratio of the intensities at E_F to the intensity of the S_1 peak

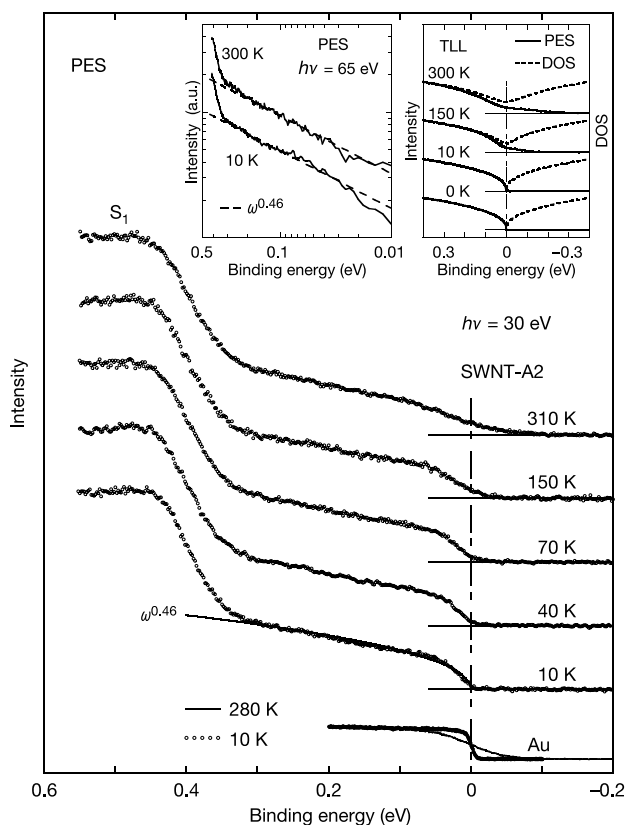


Figure 4 High-resolution photoemission spectra of the SWNT-A2 sample near E_F measured at $T = 10$ K, 40 K, 70 K, 150 K and 310 K with an energy resolution of 13 meV. These spectra were measured at HiSOR. The spectral function ω^α ($\alpha = 0.46$), broadened by the energy resolution, is indicated by a thick solid line in the spectrum at 10 K. The spectra of Au (3D conventional metal) are also shown. The left inset shows the photoemission spectra (PES), which were measured with an energy resolution of 15 meV at $h\nu = 65$ eV, plotted on a log-log scale. The right inset shows the photoemission spectra and the densities of states (DOS) calculated for the TLL state in the metallic SWNT²⁵.

plotted as a function of temperature on a log-log scale.

We consider first the following extrinsic possibilities as the origin of the peculiar spectral features near E_F . (1) An inelastic process. In this measurement, we used high-resistivity 'bucky paper'—that is, SWNTs in mat form (Methods). If the inelastic energy loss process pointed out in ref. 15 occurs in the photoemission of SWNTs, such a process should bring about a spectral profile that depends on the resistivity of the sample. As shown in the left inset of Fig. 4, however, the spectra observed at different temperatures have the same slope on a log-log scale. This indicates that an inelastic process can be excluded. (2) Alloying with Cu substrate. 'Bucky paper' with a thickness of 0.03 mm was placed on a Cu substrate. As photoemission spectroscopy has a short photoelectron mean free path (~ 0.5 nm), the electronic states of SWNTs located on the surface of the 'bucky paper' are probed. The SWNTs on the surface are not directly in contact with the substrate. Thus, the alloying effect with the Cu substrate can be neglected. (3) Broadening of VHS structures. We will consider whether the broadening of the VHS structures produces a decaying structure near E_F . The S_1 structure was well reproduced by the calculation taking into account the diameter distribution (Fig. 3). Judging from the diameter distribution of the sample, semiconducting SWNTs have zero density of states below 0.35 eV. In addition, the S_1 peak coming from semiconducting SWNTs was observed at the same energy position in the spectra measured at different temperatures (Fig. 4). Considering the temperature-dependent resistivity of the sample, a charge-up effect can be excluded. Therefore, the VHS structures do not affect the spectral profile, at least within 0.3 eV from E_F .

Another important possibility is the suppression in density of states near E_F seen in the TLL state. The density of states of the TLL state in the metallic SWNT is given as $\rho(\omega) \propto |\omega|^\alpha$ where α depends on the interaction and is written in terms of the Luttinger parameter, g , as $\alpha = (g + g^{-1} - 2)/8$ at the bulk position^{4–6}. The right inset of Fig. 4 shows the photoemission spectra calculated for the TLL state in the metallic SWNT, which were obtained on the basis of the bosonization theory of a 1D electron system²⁵. Here, $g = 0.2$ was used. The calculated spectrum has a spectral function of $\omega^{0.4}$ at zero temperature; the intensities at E_F show a power-law dependence on temperature ($\rho(E_F) \propto T^{0.4}$).

The observed spectral features are not contradictory to the theoretical calculation of the TLL state (Fig. 4). The left inset of Fig. 4 indicates that the spectra have a power-law dependence on binding energy. From fitting with ω^α , the exponent α was found to be 0.46 ± 0.10 . The spectral function obtained by broadening $\omega^{0.46}$ by the energy resolution function is represented by a thick solid line

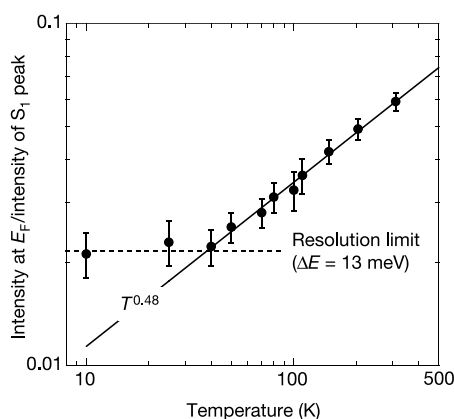


Figure 5 Temperature dependence of the ratio of the photoemission intensity at E_F to the intensity of the S_1 peak, plotted on a log-log scale. The solid line indicates the function $T^{0.48}$. The resolution limit (a horizontal dashed line) represents the intensity at E_F in the spectral function obtained by broadening $\omega^{0.46}$ by the energy resolution function with $\Delta E = 13$ meV.

in Fig. 4. The decrease in intensity near E_F was well reproduced by the calculated spectrum. The intensities at E_F are proportional to T^α ($\alpha = 0.48 \pm 0.08$), although the deviation from $T^{0.48}$ appears at temperatures below 40 K because of the broadening by an energy resolution of 13 meV (Fig. 5). An important point to note is that almost the same exponent α as the spectral function was obtained within the experimental accuracy. A value of $\alpha = 0.46$ leads to $g = 0.18$. The Luttinger parameter $g \approx 0.18$ has been obtained⁶ theoretically for a SWNT of length $3 \mu\text{m}$ with an index (10, 10); the diameter is nearly equal to the mean diameter of the present SWNT-A2 sample. This theoretical estimate is in good agreement with the present experimental one.

As mentioned above, the present photoemission spectra of the SWNT bundles show TLL behaviour. In SWNT bundles, similar TLL behaviour was also observed in the conductance measurements⁷. However, it is not obvious that the SWNTs in bundles are in isolated states without intertube interaction. We must consider whether the present spectra show the electronic states of isolated SWNTs. When metallic armchair SWNTs with the same chiral index are arranged to form a crystal-like structure in a bundle, theoretical study showed that the intertube interaction opens up a pseudogap²⁶. Indeed, Ouyang *et al.* observed the pseudogap around E_F for armchair SWNTs packed in bundles; on the other hand, such gaps were not observed for isolated armchair SWNTs²⁷. The pseudogap was interpreted to be produced by mixing between the π and π^* bands. As the present photoemission spectrum was shifted towards a higher binding energy by 0.1 eV, such a pseudogap should be observed at 0.1 eV. The present spectra did not show a gap structure (Fig. 4). We estimated the fraction of armchair SWNTs in the sample to be about 0.05. Using the diameter distribution shown in Fig. 3 inset, for example, the number of armchair SWNTs in a bundle consisting of 50 SWNTs was estimated to be only two or three. The rate at which armchair SWNTs with the same chiral index adjoin is too low for the pseudogap to be observed clearly.

The present sample consists of SWNTs with different diameters and chiralities, and is a mixture of metallic and semiconducting SWNTs. In the electronic structure calculation of such bundles, Maarouf *et al.* pointed out that the momentum mismatch between the Fermi points of neighbouring SWNTs largely suppresses the tunnelling between SWNTs²⁸. According to their calculation, even for a bundle composed of only metallic SWNTs with a random distribution of diameters and chiralities, the band structure shows no significant change near E_F (ref. 28). In the present sample, about two-thirds of the SWNTs are semiconducting, and the diameters and chiralities are inhomogeneous. The tunnelling between SWNTs will be strongly suppressed. Thus, it can be understood that the present photoemission spectra show the electronic states of isolated SWNTs as a consequence of the negligible intertube interaction suppressed by the inhomogeneity in diameter and chirality.

We now consider the influence of inhomogeneity in diameter, length and chirality of SWNTs on the exponent α . According to the theoretical calculation²⁹, the low-energy properties of metallic SWNTs do not depend on its chirality but are determined by the radius (R) and the length (L) of the tube: roughly speaking, the interaction parameter, g , is proportional to $[\ln(L/R)]^{-1/2}$. Considering the weak dependence of g on R and L , the influence of the diameter and length distribution on α can safely be neglected.

The power-law behaviours of the spectral function and intensity at E_F show good agreement with the theoretical calculation where the long-range Coulomb interaction is taken into account^{4–6}, and are consistent with the results obtained from previous transport experiments^{7–9}. In those transport measurements^{7–9}, a Coulomb blockade smeared out the power-law behaviour of the conductance at temperatures below ~ 120 K. On the other hand, photoemission measurements do not have such a disadvantage. We have found direct evidence that a TLL state is realized in SWNTs at temperatures ranging from 310 K down to at least 40 K. At temperatures below

40 K, high-resolution measurements with an energy resolution of several meV are required. The larger value of α (0.46–0.48) strongly indicates that the long-range Coulomb interaction plays an important role in SWNTs. □

Methods

Sample preparation

In photoemission measurements, a relatively large sampling area is generally needed to obtain reliable data. According to theoretical band structure calculations¹⁸, the electronic structure shows drastic changes depending on the diameter and helical arrangement of the SWNTs. It is very important to prepare a large sample consisting of SWNTs with uniform diameters. At present, however, no methods have been found that make diameters completely uniform. In this study, the SWNT samples were prepared by the laser vaporization method, in which the diameters of SWNTs can be controlled to some extent by the temperature of the furnace²². The SWNT-A (A1 and A2) and SWNT-B samples in this work were prepared using a NiCo catalyst at 1,200 °C and 1,050 °C, respectively, in the furnace. The raw soot containing SWNTs was heated to 650 °C in a dynamic vacuum to remove the fullerenes, and then refluxed in 15% H₂O₂/water solution at 100 °C for 3 h. In the reflux process, active oxygen can burn out amorphous-carbon particles at 100 °C. After washing out the brownish by-products by careful filtration, the purified SWNTs were formed into thin black paper. Finally, the remaining metal particles were vaporized in a high vacuum at 1,250 °C. The obtained 'bucky paper' had a thickness of about 0.03 mm. The lengths of the SWNTs were several micrometres. The purity of the SWNT samples was estimated by electron-energy-loss spectroscopy to be higher than 90% (ref. 30). The mean diameters of the SWNTs in the SWNT-A and SWNT-B samples were estimated to be 1.37 nm and 1.25 nm, respectively, from the A_{1g} breathing mode frequencies in the Raman spectra. The resistivity of ~18 mΩ cm at 300 K increased with decreasing temperature, and amounted to ~65 mΩ cm at 8 K. The temperature dependence was consistent with the results of ref. 31.

Measurements

The photoemission spectra were measured using a SIENTA SES200 analyser at beamlines BL-11D of the Photon Factory, High Energy Accelerator Research Organization (KEK-PF) and BL-1 of the Hiroshima Synchrotron Radiation Center (HiSOR), Hiroshima University. The total energy resolution was set to 50 meV at KEK-PF, and 13 meV or 15 meV at HiSOR. The SWNT samples were placed on a Cu substrate, because the photoemission spectra were similar as a whole to those measured using an Au substrate. The SWNT samples were cooled to 10 K. The surfaces of the samples were cleaned by heating them up to 220 °C in an ultrahigh vacuum of about 2×10^{-10} torr. The cleanliness of the sample surface was checked by the signal from the O 2p level. The zero of the binding energy was calibrated for the Fermi edge of the evaporated Au thin film. The optical absorption spectrum was measured using a HITACHI U-3500 spectrophotometer for thin films of SWNTs prepared by a splay method on a glass substrate²³. The Raman spectra were measured using a 1 m double monochromator, JOBON YVON U-1000²³. In the TEM observation, a JEOL JEM-2010F with an acceleration voltage of 120 kV was used.

Fitting procedure

The spectrum was calculated by smearing out the density of states $D(E)$ with a gaussian curve ($2\sigma = 50$ meV) as the energy resolution function and with a lorentzian curve as the spectral lifetime broadening. In the calculation, Γ_d , Γ and E_s were taken as adjustable parameters, in which the deviation was adjusted by reference to $\Gamma_d = 0.13$ nm obtained in an optical measurement²⁴, and $\Gamma_d = 0.06$ nm obtained in X-ray diffraction measurement³².

Received 20 May; accepted 22 September 2003; doi:10.1038/nature02074.

1. Voit, J. One-dimensional Fermi liquids. *Rep. Prog. Phys.* **57**, 977–1116 (1994).
2. Grioni, M. & Voit, J. in *Electronic Spectroscopies Applied to Low-dimensional Materials* (eds Hughes, H. P. & Starnberg, H. I.) 209–281 (Kluwer Academic, Dordrecht, 2000).
3. Mattsson, A. E., Eggert, S. & Johansson, H. Properties of a Luttinger liquid with boundaries at finite temperature and size. *Phys. Rev. B* **56**, 15615–15628 (1997).
4. Egger, R. & Gogolin, A. O. Effective low-energy theory for correlated carbon nanotubes. *Phys. Rev. Lett.* **79**, 5082–5085 (1997).
5. Kane, C., Balents, L. & Fisher, M. P. A. Coulomb interactions and mesoscopic effects in carbon nanotubes. *Phys. Rev. Lett.* **79**, 5086–5089 (1997).
6. Egger, R. & Gogolin, A. O. Correlated transport and non-Fermi-liquid behavior in single-wall carbon nanotubes. *Eur. Phys. J. B* **3**, 281–300 (1998).
7. Bockrath, M. *et al.* Luttinger-liquid behaviour in carbon nanotubes. *Nature* **397**, 598–601 (1999).
8. Yao, Z., Postma, H. W. C., Balents, L. & Dekker, C. Carbon nanotube intramolecular junctions. *Nature* **402**, 273–276 (1999).
9. Postma, H. W. C., de Jonge, M., Yao, Z. & Dekker, C. Electrical transport through carbon nanotube junctions created by mechanical manipulation. *Phys. Rev. B* **62**, R10653–R10656 (2000).
10. Sonin, E. B. Tunneling into 1D and quasi-1D conductors: Luttinger-liquid behavior and effects of environment. *Physica E* **18**, 331–332 (2003).
11. Sekiyama, A. *et al.* Fermi-liquid versus Luttinger-liquid behavior and metal-insulator transition in N,N'-dicyanoquinonediimine-Cu salt studied by photoemission. *Phys. Rev. B* **51**, 13899–13902 (1995).
12. Gweon, G.-H. *et al.* Fermi surfaces and single-particle spectral functions of low-dimensional inorganic non-cuprate compounds: The molybdenum bronzes. *J. Phys. Condens. Matter* **8**, 9923–9938 (1996).
13. Zwick, F. *et al.* Band mapping and quasiparticle suppression in the one-dimensional organic conductor TTF-TCNQ. *Phys. Rev. Lett.* **81**, 2974–2977 (1998).

14. Denlinger, J. D. *et al.* Non-Fermi-liquid single particle line shape of the quasi-one-dimensional non-CDW metal Li_{0.9}Mo₆O₁₇: Comparison to the Luttinger liquid. *Phys. Rev. Lett.* **82**, 2540–2543 (1999).
15. Joyn, R. Pseudogaps and extrinsic losses in photoemission experiments on poorly conducting solids. *Science* **284**, 777–779 (1999).
16. Segovia, P., Purdie, D., Hengsberger, M. & Baer, Y. Observation of spin and charge collective modes in one-dimensional metallic chains. *Nature* **402**, 504–507 (1999).
17. Voit, J. *et al.* Electronic structure of solids with competing periodic potentials. *Science* **290**, 501–503 (2000).
18. Saito, R., Dresselhaus, G. & Dresselhaus, M. S. *Physical Properties of Carbon Nanotubes* (Imperial College Press, London, 1998).
19. Wildöer, J. W. G. *et al.* Electronic structure of atomically resolved carbon nanotubes. *Nature* **391**, 59–62 (1998).
20. Suzuki, S. *et al.* Electronic structure at carbon nanotube tips studied by photoemission spectroscopy. *Phys. Rev. B* **63**, 2454181–2454187 (2001).
21. Chen, P. *et al.* Electronic structure and optical limiting behavior of carbon nanotubes. *Phys. Rev. Lett.* **82**, 2548–2551 (1999).
22. Wakabayashi, T. *et al.* Towards the selective formation of specific isomers of fullerenes: T- and p-dependence of the yield of various isomers of fullerenes C₆₀-C₈₄. *Z. Phys. D* **40**, 414–417 (1997).
23. Kataura, H. *et al.* Optical properties of single-wall carbon nanotubes. *Synth. Met.* **103**, 2555–2558 (1999).
24. Ichida, M. *et al.* Exciton effects of optical transitions in single-wall carbon nanotubes. *J. Phys. Soc. Jpn* **68**, 3131–3133 (1999).
25. Yoshioka, H. Tomonaga-Luttinger-liquid behavior in conducting carbon nanotubes with open ends. *Physica E* **18**, 212–213 (2003).
26. Kwon, Y.-K., Saito, S. & Tomanek, D. Effect of intertube coupling on the electronic structure of carbon nanotube ropes. *Phys. Rev. B* **58**, 13314–13317 (1998).
27. Ouyang, M., Huang, J.-L., Cheung, C. L. & Lieber, C. M. Energy gaps in "metallic" single-walled carbon nanotubes. *Science* **292**, 702–705 (2001).
28. Maarouf, A. A., Kane, C. L. & Mele, E. J. Electronic structure of carbon nanotube ropes. *Phys. Rev. B* **61**, 11156–11165 (2000).
29. Odintsov, A. A. & Yoshioka, H. Universality of electron correlations in conducting carbon nanotubes. *Phys. Rev. B* **59**, R10457–R10460 (1999).
30. Liu, X. *et al.* Filling factors, structural, and electronic properties of C₆₀ molecules in single-wall carbon nanotubes. *Phys. Rev. B* **65**, 0454191–0454196 (2002).
31. Vavro, J. *et al.* Electrical and thermal properties of C₆₀-filled single-wall carbon nanotubes. *Appl. Phys. Lett.* **80**, 1450–1452 (2002).
32. Abe, M. *et al.* Structural transformation from single-wall to double-wall carbon nanotube bundles. *Phys. Rev. B* **68**, 041405R (2003).

Acknowledgements We thank Y. Misaki for support with TEM observations. This study was performed with the approval of the Photon Factory Advisory Committee and under the Cooperative Research Program of HiSOR, Hiroshima Synchrotron Radiation Center, Hiroshima University. This study was supported in part by a Grant-in-Aid for Science Research from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to H.K. (kataura@phys.metro-u.ac.jp).

Migration of seismic scatterers associated with the 1993 Parkfield aseismic transient event

Fenglin Niu^{1,2}, Paul G. Silver², Robert M. Nadeau³ & Thomas V. McEvilly³

¹Department of Earth Science, MS-126, Rice University, 6100 Main Street, Houston, Texas 77005, USA

²Department of Terrestrial Magnetism, Carnegie Institution of Washington, 5241 Broad Branch Road, NW, Washington DC 20015, USA

³Berkeley Seismological Laboratory, University of California Berkeley, and Earth Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA

The time-varying deformation field within a fault zone, particularly at depths where earthquakes occur, is important for understanding fault behaviour and its relation to earthquake occurrence^{1–3}. But detection of this temporal variation has been extremely difficult, although laboratory studies have long suggested that certain structural changes, such as the properties

of crustal fractures, should be seismically detectable⁴. Here we present evidence that such structural changes are indeed observable. In particular, we find a systematic temporal variation in the seismograms of repeat microearthquakes that occurred on the Parkfield segment of the San Andreas fault over the decade 1987–97. Our analysis reveals a change of the order of 10 m in the location of scatterers which plausibly lie within the fault zone at a depth of ~3 km. The motion of the scatterers is coincident, in space and time, with the onset of a well documented aseismic transient (deformation event). We speculate that this structural change is the result of a stress-induced redistribution of fluids in fluid-filled fractures caused by the transient event.

The Parkfield region of the San Andreas fault system is an ideal natural laboratory for studying structural changes in faults⁵. For more than a decade (1987–present), this region has been heavily instrumented, containing a variety of geodetic instruments as well as a network of ultrasensitive borehole seismometers. Of particular importance to the present study is the occurrence of a remarkable aseismic transient beginning in early 1993 (Fig. 1)^{6–8} that represented a 30% increase in the slip rate along the San Andreas Fault. The transient was observed by essentially all types of available geodetic instruments, by temporal variations in seismicity rate as recorded by the borehole seismometers (the High Resolution Seismic Network, HRSN)⁹, and was accompanied by the occurrence

of four earthquakes of magnitude $M = 4+$ (that is, slightly more than four but less than five) from late 1992 to late 1994 (Fig. 1). The data suggest a complex aseismic/seismic episode, involving accelerated slip starting from the vicinity of the 1966 Parkfield earthquake, and migrating to the southeast, giving rise to geodetically observed surface slip⁸. It is, to our knowledge, the best-documented aseismic transient thus far observed.

This transient thus provides a well constrained environment to test for possible tectonically induced structural changes in the seismogenic crust. To perform this test, we have utilized tight clusters of repeating microearthquakes (magnitude $M \approx 1$) that have nearly the same location and mechanism. Recent dramatic improvements in the relative locations of small seismic events have revealed that a large fraction of seismicity occurs in these tight clusters^{10,11}. The repeating events produce nearly identical seismograms when recorded at the same station, so that temporal changes in the medium can be more easily isolated from changes in the source. Clusters have been shown to be valuable in detecting post-seismic changes in the part of the seismogram known as S-wave coda, which represents scattered S-wave energy. For example, a study of the Loma Prieta earthquake of 1989 provided strong evidence, based primarily on travel times¹² and the decorrelation of the S-wave coda¹³ that temporal variations in the wavefield could be observed from these events. Although promising, the interpretation of these post-seismic studies is complicated by the

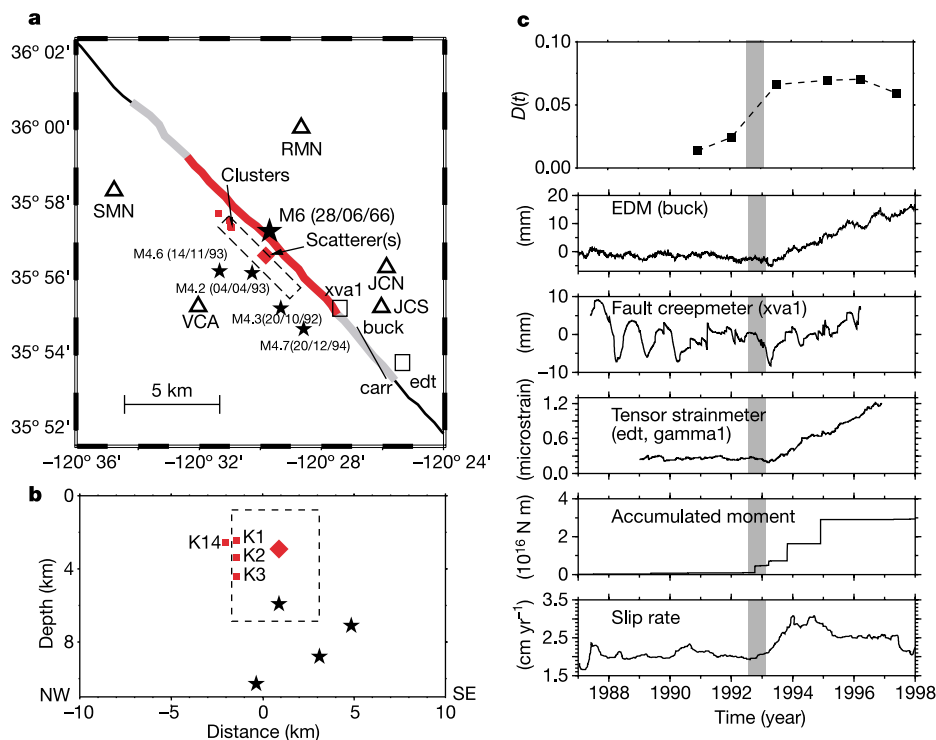


Figure 1 The Parkfield study area and geodetic results. **a**, Map view; **b**, cross-section; **c**, decorrelation index and geodetic recordings. **a**, **b**, The three-component borehole HRSN stations are shown as triangles. The red squares show the location of microearthquake clusters used in this study, denoted as K1, K2, K3 and K14. The black line shows the surface trace of the San Andreas fault. The thick grey line indicates the cross-section location shown in **b**. The large star is the epicentre of the 1966 $M = 6$ Parkfield earthquake, and the smaller stars denote the locations of the four $M = 4+$ events that occurred from 1992 to 1994. The dashed rectangle denotes the approximate location of the onset of the Parkfield aseismic transient in 1993, from recurrence intervals of event clusters⁹. The red section of fault denotes region of

increased slip from inversion of geodetic data⁸. The locations of the EDM, creepmeter and strainmeter data shown in **c** are denoted by a black line and squares, respectively. **c**, Average temporal change in decorrelation index, $D(t)$, recorded by the HRSN stations from cluster K3 in this study (see Fig. 3 and text), along with representative geodetic recordings of the Parkfield aseismic transient (EDM measurements, creepmeter, tensor strainmeters), accumulated seismic moment release, and slip rate inferred from the repeat interval of clusters of events¹⁰. The grey bar denotes the roughly six-month time interval over which the observed change in seismic structure must have occurred, based on the analysis of all clusters.

possibility that the temporal variations are due to shallow coseismic crack formation and subsequent healing near the stations, and may be unrelated to post-seismic tectonic deformation¹². In this study, we have made use of an aseismic transient to minimize the influence of coseismic crack formation. Also, the borehole HRSN has a greatly reduced environmental influence, compared to surface instruments, and thus provides an improved detection capability for subtle changes in the seismic wavefield, compared to previous studies.

We have examined a suite of candidate clusters that occurred near the $M = 6$, 1966 Parkfield earthquake (Fig. 1). The cluster

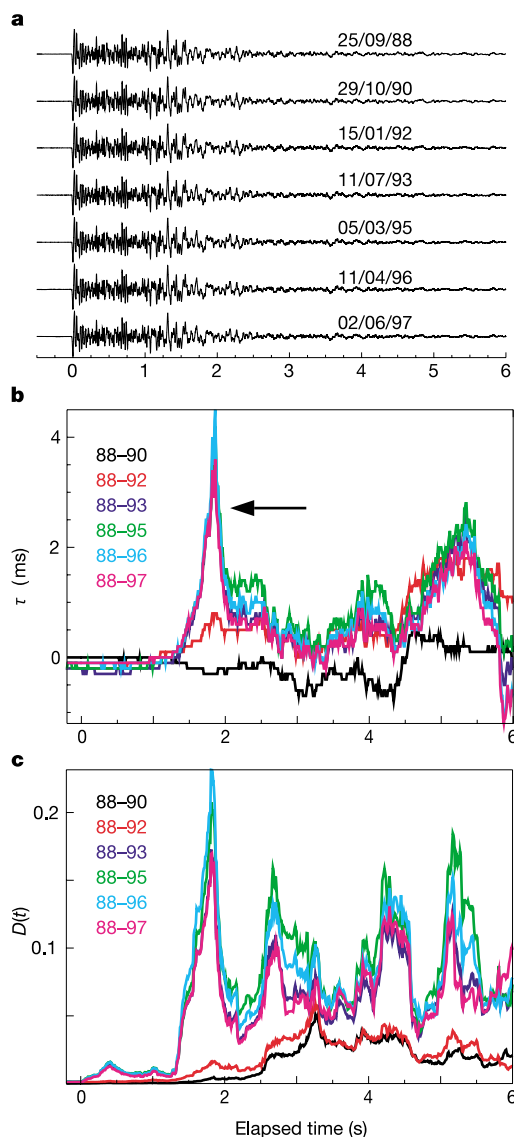


Figure 2 Cluster K3. Waveforms (a) and the calculated $\tau(t)$ (b) and $D(t)$ (c) of the components of cluster K3¹⁷ recorded at station VCA. Data are aligned to the direct P wave. Note the high waveform similarity throughout the entire seismogram. Depth of the cluster is 4 km. To quantify the difference between seismograms, we compute the cross-correlation between the first seismogram (1988) and each subsequent seismogram within a 0.5-s moving time window. The lag time $\tau(t)$ is obtained when the maximum cross-correlation, $C_m(t)$, is reached, and a decorrelation index $D(t)$ is defined as $1 - C_m(t)$. Examples of $\tau(t)$ (b) and $D(t)$ (c) observed at station VCA. Note that there is a large spike in $\tau(t)$ at $t = 1.8$ s and large changes in $D(t)$, but only after 1993.

with the greatest similarity between seismograms is referred to as K3, which is at 4 km depth; the events from this cluster produce virtually identical seismograms at a given station (Fig. 2a). Yet there are subtle differences in these records, and we seek to quantify these changes and then assess their cause. We have defined two parameters, the lag $\tau(t)$ and the decorrelation index $D(t)$ (see Methods), to quantify the difference between two seismograms. We found an obvious change that begins with the seismogram from the event occurring on 7 November 1993 (Fig. 2b, c). Of particular interest is the 'spike' of 4 ms in $\tau(t)$ about 2 s into the seismogram, which provides an important diagnostic for isolating the cause of this change. These waveform changes are clearly and systematically observed at all stations during the same time period (Fig. 3a).

We analysed a total of nine clusters located in nearly the same area as K3, so as to maintain a similar source-scatterer-receiver geometry. Among the nine clusters, we are able to confirm the observed change from the four clusters with the highest source similarity. The source similarity of a cluster is estimated by the average value, $\langle D \rangle$, of $D(t)$, for all pairs of seismograms in a cluster. Lower $\langle D \rangle$ corresponds to higher source similarity. The $\langle D \rangle$ values are 0.03–0.07 for the four clusters where the signal is observed, and 0.10–0.18 for the remaining five clusters. In each of the four clusters, the subtle change is relatively easy to identify from the two best stations, VCA and RMN (based on a comparison of station-specific source similarity). By combining the observations of different clusters, we can confine the time interval over which the medium change could have occurred to ~ 6 months (between 14 July 1992 and 5 February 1993; Fig. 1c).

A subtle but systematic difference in the waveforms in a cluster

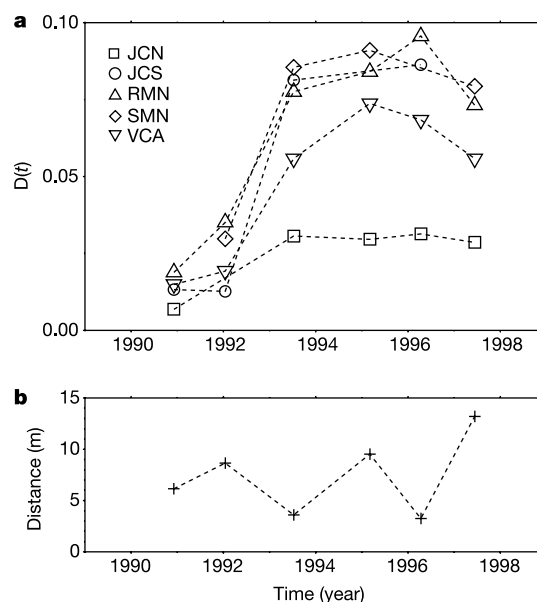


Figure 3 Temporal change in average $D(t)$ and relative distances of events with respect to the first event. **a**, Note the remarkable increase in $D(t)$ (averages calculated over 6 s of elapsed time) between 1992 and 1993 for all stations. The step-function increase in $D(t)$ suggests a temporal change in the medium at this time. **b**, Temporal changes of $\tau(t)$ and $D(t)$ could be due to systematic differences in earthquake location, if, for example, the pre- and post-1993 events are located in two separate subclusters. To test this, we measured the S-P travel times of each event-station pair and relocated all subsequent events relative to the first event. If the observed change in $D(t)$ is due to systematic variations in event location, these plots should have the same form as **a**. We see no such pattern, and conclude that the temporal variations in earthquake location did not produce the systematic variation in $D(t)$.

is most probably due to a change in the medium, or a slight difference of source parameters, or both. Numerical experiments, however, suggest that both background-velocity and earthquake-location models produce only long-period changes (Fig. 4) instead of the 'spike' variation shown in Fig. 2. On the other hand, a change in the location of a single scatterer results in a high-frequency and isolated 'spike' in $\tau(t)$ and $D(t)$ (Fig. 4). The same result could be caused by a localized reduction in velocity near the scatterer. The striking similarity between this synthetic and the lag time series in Fig. 2b and c for VCA strongly suggests that we are observing a change in the location of a single scatterer, or possibly a localized change in velocity in the vicinity of the scatterer.

In addition, if earthquake location were the cause of the systematic pattern in Figs 2 and 3a, the pre- and post-1993 events would have to be grouped into two separate subclusters. This is unlikely for two reasons. First, we measured the S-P times and relocated all subsequent events with respect to the initial event. If there were indeed two subclusters, then the distance from these events to the initial event should have the same form as Fig. 3a. In fact, the estimated distance pattern for K3 (Fig. 3b) shows no resemblance to the observed pattern in $D(t)$ (Fig. 3a). Second, such an explanation would require that all four clusters are coincidentally composed of two subclusters: the pre- and post-1993 events. From these considerations we conclude that the observed temporal variation is not

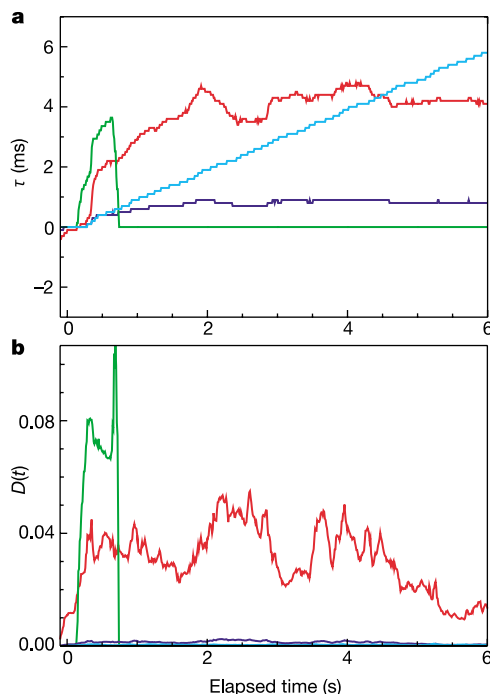


Figure 4 Quantities derived from synthetic seismograms **a**, $\tau(t)$; **b**, $D(t)$. Synthetics are calculated using an acoustic medium with a random distribution of scatterers, illustrating the effect of various temporal changes on these two functions: event location change of 2 m (blue) and 10 m (red), a decrease in background velocity of 0.1% (turquoise), and a 10-m change in the location of a single scatterer. All of these sources of change produce long-period variations, except a single scatterer, which produces a 'spike'. The observed spike (Fig. 2b) and other excursions are thus most easily explained by a change in the location of one, or only a few, scatterers. Although the acoustic synthetics is not strictly applicable to the problem of S waves, which requires a fully elastic treatment¹⁸, it nevertheless fits our purpose of illustrating different effects due to changes in source and media rather than modelling the exact coda waveforms.

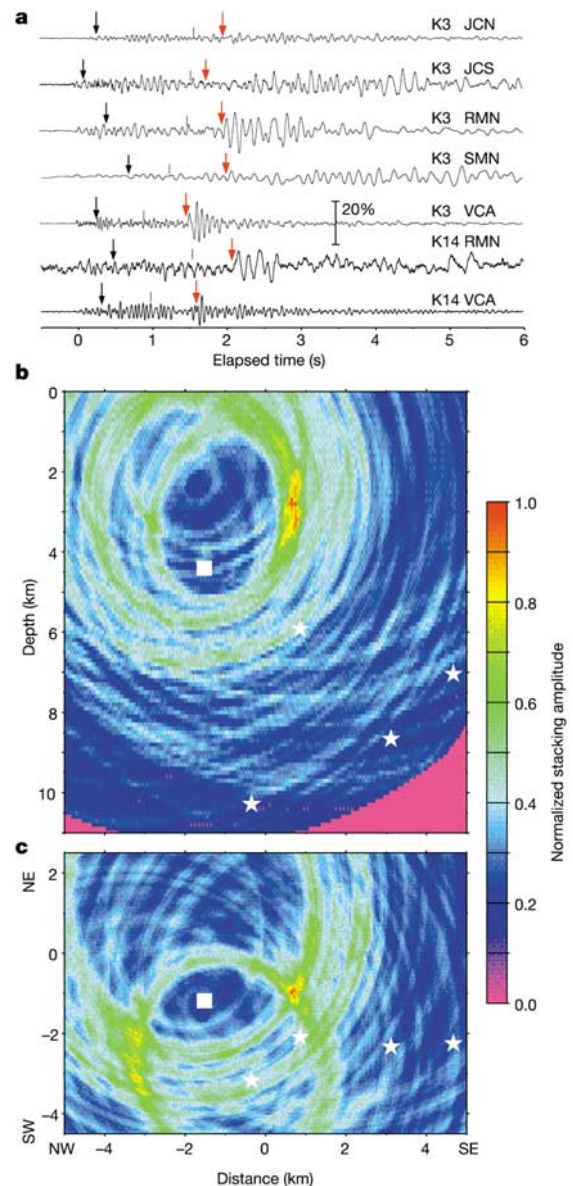


Figure 5 Differential vertical-component seismograms and the locations of the 'moving' scatterer(s). **a**, Differential vertical-component seismograms, $\delta s(t)$. Difference is taken between the average of post- and pre-1993 records. These seismograms enhance the parts of the seismogram that have changed in time. The records for VCA and RMN clearly show the contribution from a time-variable scatterer. Vertical bar denotes 20% of original P-wave amplitude. Note that for VCA, differential amplitude is comparable to overall coda amplitude in Fig. 2a. These seismograms were used to locate the time-variable scatterer by migration. Shown is the distribution of stacked amplitude, $A(\mathbf{x})$, as a function of position $\mathbf{x}$. It is normalized to the highest value. **b**, **c**, Cross-section parallel to the fault (**b**) and a map view at 3 km depth (**c**) that passes through the highest value of $A(\mathbf{x})$. The map reveals a well defined maximum (red spot) that is localized approximately 2 km southeast of the cluster along the San Andreas fault and at 3 km depth. Also shown is the location of cluster K3 (white square) and the locations of the four $M = 4+$ events that occurred between 1992 and 1994 (white stars). Note that the scatterer is closest to the event occurring on 4 April 1993. The medium change precedes this event. Predicted S-wave scatterer arrival times are shown by red arrows in (**a**). Also shown are the predicted direct-S arrival (vertical lines) and the predicted P-wave arrival times from the same scatterer (black arrows). Note that there is no comparable signal in the P-coda, consistent with the hypothesis that the scatterer is due to a fluid-filled fracture, which more efficiently reflects S than P.

due to differences in earthquake location, and is instead due to a change in the medium.

We have utilized a simple procedure for locating spatially the scatterers whose properties have changed in time, using differential seismograms (Fig. 5a) (see Methods). Stacking of differential seismograms clearly suggests a 'moving scatterer' located about 2 km to the southeast of the cluster (Fig. 5b) and at approximately 3 km depth (Fig. 5c). It corresponds to the 'spike' in lag time shown in Fig. 2b. Here we used a constant-velocity path model for each station such that the S-arrival times are fitted exactly. Such an approach approximately accounts for the three-dimensional structure where the direct-S and scattered-S paths are similar, and thus corrects for near-source structure as well as the particularly strong shallow heterogeneity beneath the receiver. There remains a possible error in scatterer location from three-dimensional structure that we estimate to be a few hundred metres, but which has essentially no effect on our scientific conclusions.

We also regard surface environmental effects, such as precipitation, to be an unlikely cause of the change in the medium. The largest effect related to precipitation will be an annual cycle, which would have little effect on the pattern in Fig. 3. Although 1993 did have increased precipitation compared to earlier years, the highest precipitation was in 1995, and this had no discernible effect on our data, as is clear from Fig. 3. Perhaps the strongest argument against an environmental explanation is that the depth of the resolved scatterer is large (3 km) and far removed from surface environmental effects.

From the high amplitude of the coda, we assume that the system is in the strong scattering regime, where the wavelength of the interacting wave is comparable to the size of the scatterer¹⁴. From this and the characteristic frequency of the seismic waves (10 Hz) we estimate the dimension of the scatterer to be of the order of 300 m. Assuming that scatterer migration is the cause of the time variation, we estimate a scatterer displacement of the order of 10 m from the 4-ms 'spike' (Fig. 2b; see Methods). Given the large magnitude of displacement, it is highly unlikely that this is the motion of a solid scatterer embedded within the medium; instead, it suggests that the scatterer is a fluid-filled fracture (or system of fractures) with the appropriate dimensions, containing fluid whose centroid has migrated 10 m. Such a fracture effectively represents a free-slip boundary and is a very efficient scatterer for shear waves polarized parallel to the plane of the fracture surface, although much less so for P waves. This may account for the absence of a comparable signal in the P-wave coda from this scatterer(s) (Fig. 5). If, alternatively, the observed change of waveform is caused by a velocity decrease near the scatterer (as mentioned above), it would most probably be due to a localized change in the density of fluid-filled cracks. Such a change would be equivalent to scatterer migration from a physical viewpoint, because both explanations appeal to fluid migration at or near the scatterer.

One possible explanation for this structural change is the coseismic deformation associated with the first of the four $M = 4+$ events (20 October 1992). In general, the primary influence of coseismic deformation is approximately within a single fault dimension, ~ 1 km for an $M = 4+$ earthquake¹⁵. The scatterer is, however, about 6 km (six fault lengths) away from the 20 October 1992 earthquake, and probably too far to have a significant effect, although we cannot entirely exclude such a linkage. On the other hand, given the close correspondence (both spatially and temporally) between the change in the medium and the onset of the 1993 transient (Fig. 1), it is highly likely that this structural change was a further manifestation of the aseismic transient. Indeed, the inferred location of the scatterer essentially coincides with the zone of accelerated slip⁹ that was active during the observed time of the scatterer change (Fig. 1). We thus speculate that the inferred fluid migration is either the direct result of the accelerated slip, or a more general response to the stress perturbation that gave rise to the transient event. $\square$

Methods

Calculation of lag $\tau(t)$ and decorrelation index $D(t)$

To quantify the difference between seismograms in a cluster, we compute the cross-correlation between the first seismogram and each subsequent seismogram within a 0.5-s moving time window. The lag time $\tau(t)$ is obtained when the maximum cross-correlation, $C_m(t)$, is reached, and a decorrelation index $D(t)$ is defined as $1 - C_m(t)$. Both parameters constitute statistical measures of scatterer lag times. For a random distribution of scatterers and for characteristic frequency ω , $\tau(t)$ can be interpreted as the weighted mean $\langle \tau_i \rangle$ and $2D/\omega^2$ as the weighted variance σ_τ^2 (weighted by squared amplitude) of the time lags, τ_i , associated with the individual scatterers in the time window¹⁶. In relevant special cases, an actual time shift, t_s , can be retrieved from these statistics. For example, consider two populations of scatterers in a time window that are either shifted by t_s or unshifted. It can then be shown that $t_s = (\sigma_\tau^2 + \langle \tau_i \rangle^2) / \langle \tau_i \rangle$. For a single shifted scatterer in the time window, $t_s = \langle \tau_i \rangle$. In the more general case, $\langle \tau_i \rangle$ underestimates t_s . For our data, with a 20-Hz characteristic frequency, $\sigma_\tau \approx \langle \tau_i \rangle$, in which case we estimate the bias to be a factor of 2.

Locating 'moving' scatterers

To locate spatially the scatterer(s) whose properties have changed in time, we first constructed a differential seismogram $\delta s(t)$ by taking the difference of the average pre- and post-1993 seismograms for each station. For noise-free seismograms, $\delta s(t)$ will consist only of energy from any scatterer whose physical characteristics (location of scatterer, localized velocity change near the scatterer, or scatterer strength) have changed the amplitude or travel time of the scattered wave. Energy from an unchanged scatterer will be removed by this procedure. We then computed an n th-root ($n = 1$) summed amplitude, $A(x)$, of $\delta s(t)$ from each station, stacking on the predicted arrival time of a candidate scatterer originating from location x . The location of the scatterer is determined as the maximum of the amplitude function, $A(x)$.

Received 1 November 2002; accepted 15 October 2003; doi:10.1038/nature02151.

- Stein, R. S. The role of stress transfer in earthquake occurrence. *Nature* **402**, 605–609 (1999).
- Freed, A. M. & Lin, J. Delayed triggering of the 1999 Hector Mine earthquake by viscoelastic stress transfer. *Nature* **411**, 180–183 (2001).
- Bowman, D. D., Ouzillon, G., Sammis, C. G., Sornette, A. & Sornette, D. An observational test of the critical earthquake concept. *J. Geophys. Res.* **103**, 24359–24372 (1998).
- Nur, A. Effects of stress on velocity anisotropy in rocks with cracks. *J. Geophys. Res.* **76**, 2022–2034 (1971).
- Bakun, W. H. & Lindh, A. G. The Parkfield, California, earthquake prediction experiment. *Science* **229**, 619–624 (1985).
- Gwyther, R. L., Gladwin, M. T., Mee, M. & Hart, R. H. G. Anomalous shear strain at Parkfield during 1993–94. *Geophys. Res. Lett.* **23**, 2425–2428 (1996).
- Langbein, J., Gwyther, R. L., Hart, R. H. G. & Gladwin, M. T. Slip-rate increase at Parkfield in 1993 detected by high-precision EDM and borehole tensor strainmeters. *Geophys. Res. Lett.* **26**, 2529–2532 (1999).
- Gao, S. S., Silver, P. G. & Linde, A. T. Analysis of deformation data at Parkfield, California: Detection of a long-term strain transient. *J. Geophys. Res.* **105**, 2955–2968 (2000).
- Nadeau, R. M. & McEvilly, T. V. Fault slip rates at depth from recurrence intervals of repeating microearthquakes. *Science* **285**, 718–721 (1999).
- Nadeau, R. M., Antolik, M., Johnson, P. A., Foxall, W. & McEvilly, T. V. Seismological studies at Parkfield III: Microearthquake clusters in the study of fault-zone dynamics. *Bull. Seismol. Soc. Am.* **84**, 247–263 (1994).
- Waldhauser, F. & Ellsworth, W. L. A double-difference earthquake location algorithm: method and application to the northern Hayward fault. *Bull. Seismol. Soc. Am.* **90**, 1353–1368 (2000).
- Dodge, D. A. & Beroza, G. C. Source array analysis of coda waves near the 1989 Loma Prieta, California, mainshock: Implications for the mechanism of coseismic velocity changes. *J. Geophys. Res.* **102**, 24437–24458 (1997).
- Baisch, S. & Bokelmann, G. H. Seismic waveform attributes before and after the Loma Prieta earthquake: Scattering change near the earthquake and temporal recovery. *J. Geophys. Res.* **106**, 16323–16337 (2001).
- Aki, K. & Richards, P. G. *Quantitative Seismology: Theory and Methods* (Freeman, San Francisco, 1980).
- Kanamori, H. & Anderson, D. L. Theoretical basis of some empirical relations in seismology. *Bull. Seismol. Soc. Am.* **65**, 1073–1095 (1975).
- Snieder, R., Gert, A., Douma, H. & Scales, J. Coda wave interferometry for estimating nonlinear behavior in seismic velocity. *Science* **295**, 2253–2255 (2002).
- Nadeau, R. M. & McEvilly, T. V. Seismological studies at Parkfield; V. Characteristic microearthquake sequences as fault-zone drilling targets. *Bull. Seismol. Soc. Am.* **87**, 1463–1472 (1997).
- Poupinet, G., Ratdompurbo, A. & Coutant, O. On the use of earthquake multiplets to study fractures and the temporal evolution of an active volcano. *Ann. Geofis.* **39**, 253–264 (1996).

Acknowledgements We thank the people who are involved in the installation and maintenance of the High Resolution Seismic Network, and also thank M. Fehler, P. Malin, E. Roeloffs, C. Thurber and L. Wen for discussions, and A. Rubin and S. Roecker for reviews. This work was supported by the Carnegie Institution of Washington, Rice University, NASA and USGS. Partial processing of the data was done at the University of California's Berkeley Seismological Laboratory and at the Center for Computational Seismology (CCS) at the Lawrence Berkeley National Laboratory.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to F.N. (niu@rice.edu).

Oligocene mammals from Ethiopia and faunal exchange between Afro-Arabia and Eurasia

John Kappelman¹, D. Tab Rasmussen³, William J. Sanders⁴, Mulugeta Feseha², Thomas Bown⁵, Peter Copeland⁶, Jeff Crabaugh⁷, John Fleagle⁸, Michelle Glantz⁹, Adam Gordon¹, Bonnie Jacobs¹⁰, Murat Maga¹, Kathleen Muldoon³, Aaron Pan¹⁰, Lydia Pyne¹, Brian Richmond¹¹, Timothy Ryan¹², Erik R. Seiffert¹², Sevket Sen¹³, Lawrence Todd⁹, Michael C. Wiemann¹⁴ & Alisa Winkler^{10,15}

¹Department of Anthropology, and ²Department of Geological Sciences, University of Texas, Austin, Texas 78712, USA

³Department of Anthropology, Washington University, St Louis, Missouri 63130, USA

⁴Museum of Paleontology, University of Michigan, Ann Arbor, Michigan 48109, USA

⁵2300 Arapahoe Avenue, no. 236, Boulder, Colorado 80302, USA

⁶Department of Geosciences, University of Houston, Houston, Texas, 77204-5503, USA

⁷Department of Geology and Geophysics, University of Wyoming, Laramie, Wyoming 82071-3006, USA

⁸Department of Anatomical Sciences, State University of New York, Stony Brook, New York 11794-8081, USA

⁹Department of Anthropology, Colorado State University, Fort Collins, Colorado 80523-1787, USA

¹⁰Department of Geological Sciences, Southern Methodist University, Dallas, Texas 75275-0395, USA

¹¹Department of Anthropology, George Washington University, Washington DC 20052, USA

¹²Department of Biological Anthropology and Anatomy, Duke University, Durham, North Carolina 27708-0383, USA

¹³Museum National d'Histoire Naturelle, Laboratoire de Paléontologie, 8 rue Buffon, 75005 Paris, France

¹⁴Center for Wood Anatomy Research, USDA Forest Service, Forest Products Laboratory, One Gifford Pinchot Drive, Madison, Wisconsin 53705-2398, USA

¹⁵Department of Cell Biology, University of Texas Southwestern Medical Center, Dallas, Texas 75390, USA

Afro-Arabian mammalian communities underwent a marked transition near the Oligocene/Miocene boundary at approximately 24 million years (Myr) ago. Although it is well documented that the endemic paenungulate taxa were replaced by migrants from the Northern Hemisphere, the timing and evolutionary dynamics of this transition have long been a mystery because faunas from about 32 to 24 Myr ago are largely unknown¹. Here we report a late Oligocene fossil assemblage from Ethiopia, which constrains the migration to postdate 27 Myr ago, and yields new insight into the indigenous faunal dynamics that preceded this event. The fauna is composed of large paenungulate herbivores and reveals not only which earlier taxa persisted into the late Oligocene epoch but also demonstrates that one group, the Proboscidea, underwent a marked diversification. When Eurasian immigrants entered Afro-Arabia, a pattern of winners and losers among the endemics emerged: less diverse taxa such as *arsinoitheres* became extinct, moderately species-rich groups such as hyracoids continued into the Miocene with reduced diversity, whereas the proboscideans successfully carried their adaptive radiation out of Afro-Arabia and across the world.

The newly discovered fossils are from the Chilga region of Ethiopia's northwestern plateau at an elevation of about 1,950 m (Fig. 1). Sedimentary rocks total at least 130 m and are exposed in a series of stream and gully cuts²⁻⁵. The channel and floodplain deposits include reworked volcanoclastics, common but usually thin lignites², and claystones and siltstones that often represent

palaeosols with varying degrees of maturity. Fossils have been found in each of these settings. Earlier work with a rich pollen assemblage³ suggests that the sediments were deposited at an elevation of about 1,000 m.

The sediments conformably overlie the uppermost Ethiopian plateau flood basalts. A basalt sample from the base of the section (Fig. 2) provides a whole-rock K–Ar date of 32.4 ± 1.6 Myr (see Supplementary Information) and offers a maximum age for the sediments. Duplicate ^{40}Ar – ^{39}Ar analyses of K-feldspar⁶ from a tuff within the fossiliferous portion of the sedimentary section yield ages with a weighted average of 27.36 ± 0.11 Myr (see Supplementary Information). These data, along with the results from palaeomagnetic studies⁷, support a correlation of the Chilga section with Chron C9n (ref. 8) (see Methods and Fig. 2). Our dates are consistent with new dates for the Ethiopian plateau flood basalts in both Ethiopia⁹⁻¹¹ and Yemen¹². Previous work³ reported that the Chilga sediments were no older than late Miocene in age, but the basalt that was dated by the whole-rock K–Ar technique at 8.0 ± 1.2 Myr is not conformable with the sediments.

There are now over 70 localities at Chilga with vertebrates, invertebrates, plant macrofossils and trace fossils (burrows, termitaria and rhizoliths). Vertebrate remains are usually fragmentary and represent medium- to large-sized herbivores. A possible preservational bias against small mammals may be due to diagenetic leaching of bone before fossilization. Vertebrates and plant macrofossils are often found together, with some leaf impressions preserved in clays that drape fossil vertebrates. Such occurrences are usually rare; the Chilga finds are significant in proving that these plants and animals were contemporaneous. Details of the macroflora will be provided elsewhere.

The Chilga mammals provide evidence for the continued evolution and diversification of some of the most common elements of the endemic Fayumian fauna (Table 1). One of the most distinctive taxa is a new species of *Arsinoitherium* (order Embrithopoda), well known for its distinctive horns and high-crowned molars. This new species (Fig. 3a–c) is the largest yet known of the genus. *Arsinoitherium* is best known from Palaeogene deposits of Egypt, Libya, Oman, and Angola¹³⁻¹⁷ and its occurrence at Chilga is the last

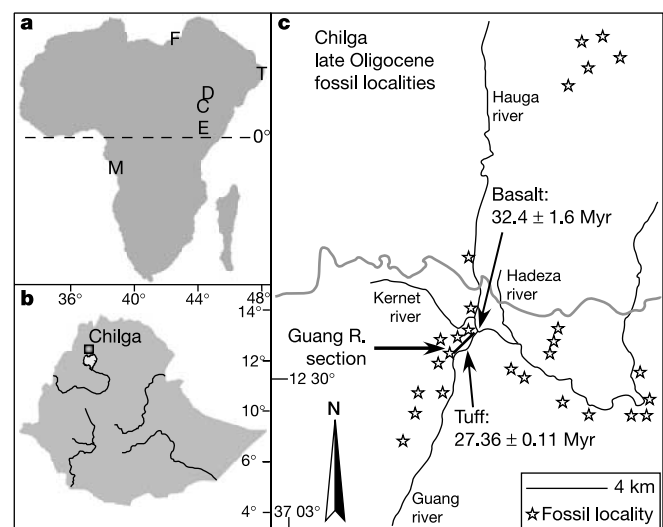


Figure 1 Map of Chilga and surrounding area containing the fossil localities. **a**, Map of Afro-Arabia with several important Palaeogene localities including Chilga (C; Ethiopia), Dogali (D; Eritrea) (ref. 27), Eragaleit (E; Kenya) (ref. 29), Fayum (F; Egypt) (ref. 13), Malembe (M; Angola) (ref. 16) and Thaytini and Taqah (T; Oman) (ref. 17). **b**, Location of Chilga in Ethiopia and north of lake Tana. **c**, Detailed map of the Chilga area showing the fossil localities, geological section (see Fig. 2) and dated rock samples (see Supplementary Information) along the Guang and Hauga rivers.

known for the order Embrithopoda.

Hyracoids are moderately diverse at Chilga and document the continued presence of two well known but conservative Fayumian taxa. *Pachyhyrax crassidentatus* and *Megalohyrax eocaenus* are the most common large hyracoids from the early Oligocene^{18,19}, with the latter first appearing in the late Eocene and changing little over its duration of several Myr. The new Chilga taxa are distinct from these at the specific level, with *Pachyhyrax* sp. nov. being smaller than *P. crassidentatus* and differing from it in its relatively higher, sharper and more delicate crests, whereas *Megalohyrax* sp. nov. (Fig. 3d) is remarkably close in size and structure to *M. eocaenus* but differs in premolar proportions. A third, larger hyracoid at Chilga is a new genus probably related to *Pachyhyrax*, whereas a fourth is closely related to *Bunohyrax*.

The Proboscidea has ancient Afro-Arabian roots^{20,21} and is the most diverse order of mammals at Chilga, represented by three families and five new species. Three of the Chilga proboscideans belong to the Palaeomastodontidae, a primitive family previously known from the late Eocene and early Oligocene of Afro-Arabia²². Two of these new species have close taxonomic ties to the genus *Palaeomastodon*, but are larger than the Fayum *Palaeomastodon* (Fig. 3e, f). The larger has some molar features that anticipate the morphology of the early Miocene mammutid *Eozygodon* (Fig. 3f). A

third new species (Fig. 3g) is closely aligned with the palaeomastodont *Phiomia*; its molars equal or exceed the uppermost size range of Fayum specimens of this genus and it has an outsized mandibular symphysis considerably longer than those of its Fayumian congeners. The Chilga palaeomastodonts represent the youngest-known record for this family.

Chilga also documents the oldest occurrence of deinotheres (Fig. 3h, i). This distinctive family of proboscideans was previously known from the early Miocene through to the early Pleistocene epoch but its origins are poorly understood²³. The new taxon clearly differs from other deinotheres, particularly in the nascent development of a third loph(id) in deciduous fourth premolars and first molars. The occlusal morphology of its cheek teeth suggests independent derivation from a bunolophodont form such as *Moeritherium*^{23,24}, rather than sharing an earlier ancestry with barytheres and numidotheres in the lophodont barytherioid group. The recovery of deinotheres at 27 Myr ago extends the temporal range of this group by 7 Myr.

A third Chilga proboscidean family, Gomphotheriidae, is represented by a rare species with molars closely resembling those of primitive *Gomphotherium* (Fig. 3j), which previously had an earliest appearance of approximately 20–18 Myr ago in Africa and Eurasia^{25–27}. A similarly ancient but undescribed proboscidean with possible affinity to the Chilga gomphotheres is known from Dogali, Eritrea (Fig. 1a)²⁸. The Chilga specimens are among the smallest to be attributed to Elephantoida and markedly extend the known duration of this family back into the late Oligocene.

The recovery of these fossils from Ethiopia enlarges our understanding of some of the ecological attributes of Afro-Arabia's Oligocene fauna. For example, arsinotheres, hyracoids, and palaeomastodonts are best known from Palaeogene sites in Egypt, Libya, Oman and Angola^{13–17} (Fig. 1), and the shared continental margin setting suggested that their habitat preferences were limited to coastal and estuarine environments. The discovery of these taxa in

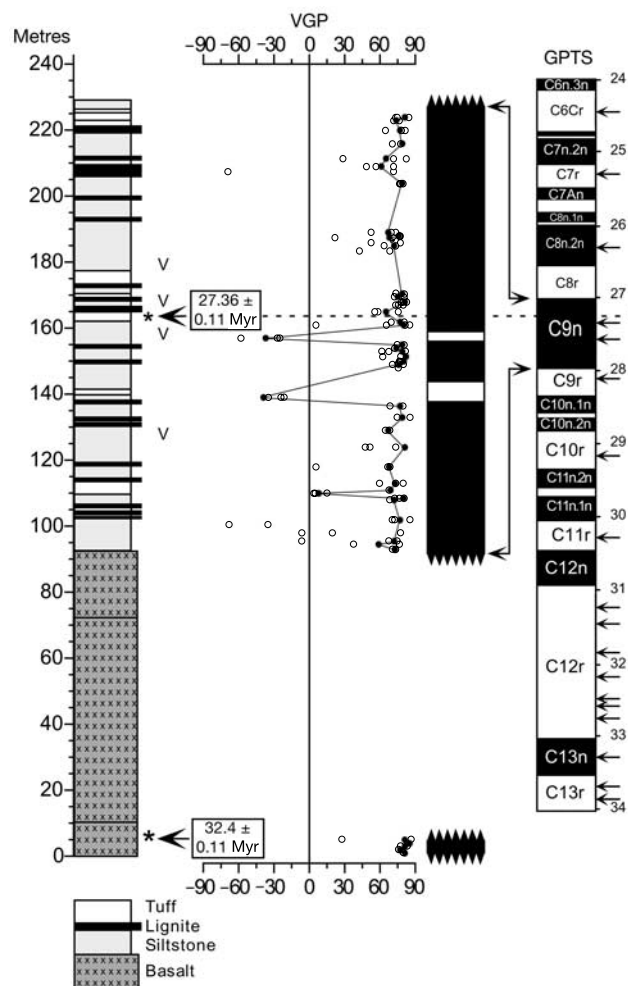


Figure 2 The Chilga section preserves volcanics and fluvial sediments. Vertebrate fossils date to around 27 Myr ago (see Methods for details). Ages along right of the GPTS panel are in Myr. V, vertebrate localities. VGP, virtual geomagnetic pole. White circles, samples; black circles, site means. Black bars, normal polarity; white bars, reversed polarity. Arrows, tiny wiggles.

Table 1 Taxonomy and temporal occurrence of Chilga mammals

Taxonomic group	Fayumian Eocene/Oligocene	Chilga late Oligocene	Rusingan early Miocene
Proboscidea			
Palaeomastodontidae	X	X	—
<i>Palaeomastodon beadnelli</i>	x	—	—
<i>Palaeomastodon parvus</i>	x	—	—
<i>Palaeomastodon intermedius</i>	x	—	—
aff. <i>Palaeomastodon</i> sp. nov. A	—	x	—
aff. <i>Palaeomastodon</i> sp. nov. B	—	x	—
<i>Phiomia wintoni</i>	x	—	—
<i>Phiomia minor</i>	x	—	—
<i>Phiomia osborni</i>	x	—	—
<i>Phiomia</i> sp. nov.	—	x	—
Gomphotheriidae	—	X	X
cf. <i>Gomphotherium</i> sp. nov.	—	x	—
<i>Gomphotherium angustidens</i>	—	—	x
Deinotheriidae	—	X	X
Gen. et sp. nov.	—	x	—
<i>Prodeinotherium hoblely</i>	—	—	x
Hyracoidea			
Saghatheriidae	X	X	—
<i>Pachyhyrax crassidentatus</i>	x	—	—
<i>Pachyhyrax</i> sp. nov.	—	x	—
Gen. et sp. nov., aff. <i>Pachyhyrax</i>	—	x	—
<i>Megalohyrax eocaenus</i>	x	—	—
<i>Megalohyrax</i> sp. nov.	—	x	—
<i>Bunohyrax fajumensis</i>	x	—	—
<i>Bunohyrax major</i>	x	—	—
<i>Bunohyrax</i> sp.	—	x	—
Titanohyracinae*	X	—	X
Embrithopoda			
Arsinoitheriidae	X	X	—
<i>Arsinoitherium zitteli</i>	x	—	—
<i>Arsinoitherium</i> sp. nov.	—	x	—

Taxonomic groups include order (in bold font), family/subfamily, genus and species. X, family or subfamily present; x, genus and species present; —, not present.

*Although this subfamily does not occur at Chilga, titanohyracines are known from either side of the transition but have much reduced diversity in the early Miocene.

the highlands of northwestern Ethiopia instead suggests that they were widespread herbivores in Afro-Arabia, and occupied generalist niches with broad ecological tolerances. This observation is especially interesting in light of their overall low specific diversities.

Before the discovery of the Chilga fauna many aspects of Afro-Arabia's Oligocene/Miocene faunal transition were matters of speculation. It is now clear that the success of the invading northern immigrants was not a consequence of their movement into an

ecological vacuum created by a much earlier extinction of the Afro-Arabian endemics; instead, the new fossils provide evidence that they evolved and diversified through the late Oligocene. Some taxa, such as arsinotheres and hyracoids, continued a conservative evolutionary trajectory with size changes and slight morphological innovation, whereas others, such as the proboscideans, underwent a radiation that resulted in the origin of new families including the gomphotheres and deinotheres. At present, there is no indication

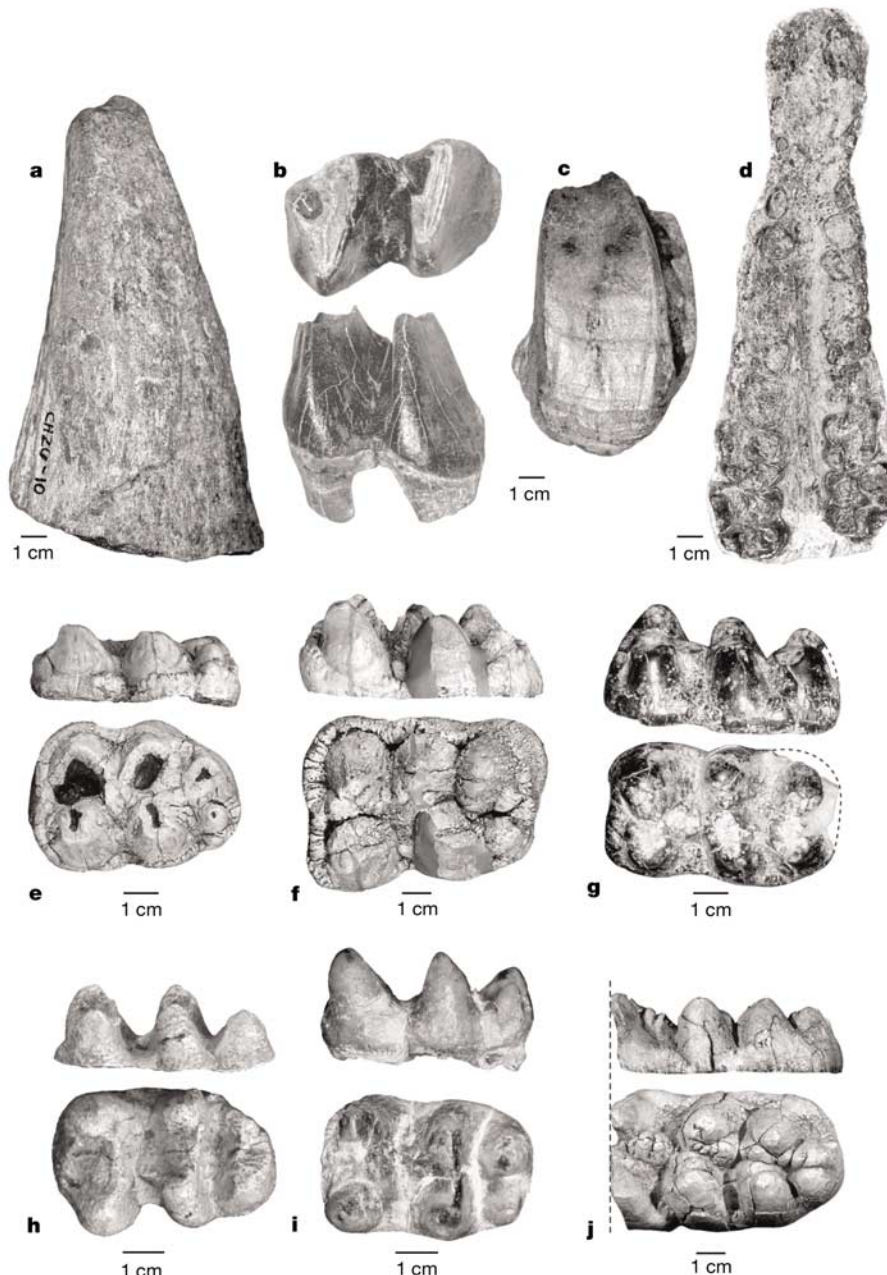


Figure 3 Fossils representing paenungulate taxa from Chilga. **a**, Specimen CH26-10. *Arsinoitherium* sp. nov., partial horn core (left side?), height = 225 mm. **b**, Specimen CH25-17. *Arsinoitherium* sp. nov., lower molar, occlusal and buccal views. **c**, Specimen CH3-95. *Arsinoitherium* sp. nov., upper molar, distal view, height = 131 mm. **d**, Specimen CH18-1. *Megalohyrax* sp. nov., palate, length = 210 mm. **e**, Specimen CH35-23. New species, affinity to *Palaeomastodon* (aff. *Palaeomastodon* sp. nov. A), right third upper molar (M^3), buccal and occlusal views. Note the incomplete trilophodonty of the specimen. **f**, Specimen CH14-11. New species, affinity to *Palaeomastodon* (aff. *Palaeomastodon* sp. nov. B), right M^3 , buccal and occlusal views. Note the incomplete trilophodonty of the specimen and its more rectangular occlusal outline in comparison

with CH35-23. **g**, Specimen CH 9-1. *Phiomia* sp. nov., left M^2 , lingual and occlusal views; portion of posterior end is missing (dotted line). **h**, Specimen CH35-3c. Deinotheriidae gen. et sp. nov., left first lower molar (M_1), buccal and occlusal views. **i**, Specimen CH35-3a. Deinotheriidae gen. et sp. nov., right fourth lower deciduous premolar (DP_4), lingual and occlusal views. **j**, Specimen CH14-14. Cf. *Gomphotherium* sp. nov., right M_3 , lingual and occlusal views; anterior end is missing (dotted line). Note the rounded conelets, transverse continuity of half-lophids, larger posterior 'heel', and trefoil arrangement of outer main conelets and anterior and posterior accessory conules on the pretrite side.

that contact with the invading migrants spurred on any of these changes; rather, fluctuations in environmental conditions possibly drove their continued evolution⁹. The discovery of new fossil localities dating to the Oligocene/Miocene boundary will be required to test the remaining possibility that competitive exclusion between the Afro-Arabian endemics and the invading immigrants was responsible for the extinction of the former. It seems likely that the phyletic conservatism and perhaps generalized habits of many of the endemic taxa greatly limited their ability to compete with the invaders; other endemics such as the proboscideans, which underwent greater diversification and specialization in the Afro-Arabian Oligocene, ended up on the winning side of the equation. □

Methods

Chronology

The Chilga section has more than 90 m of volcanics at its base that are overlain by at least 130 m of fluvial sediments. The basalt at the base of the section is dated at 32.4 ± 1.6 Myr (see Supplementary Information) and provides a maximum age for the section. Studies of the isothermal remanent magnetism⁷ for a suite of siltstones show a dominance of low-coercivity grains indicating magnetite or maghaematite, an expected result given that basalts form the dominant parent material. All sediments show some evidence of intermediate-coercivity grains indicating specular haematite as well as small amounts of high-coercivity grains suggesting pigmentary haematite and goethite, but these are both minor constituents relative to the low-coercivity fraction. Stepwise alternating-field demagnetization was carried out on 118 samples with generally three samples per stratigraphic level. Palaeomagnetic reversal stratigraphy demonstrates a dominance of normal polarity. Duplicate ⁴⁰Ar–³⁹Ar age spectra analyses of K-feldspar each separated from a tuff at 165 m produced a weighted average of 27.36 ± 0.11 Myr (see Supplementary Information) that provides an absolute tie point to Chron C9n ($27.004\text{--}27.946$ Myr). Small arrows to the far right of the geomagnetic polarity timescale (GPTS)⁸ in Fig. 2 represent “tiny wiggles” and the two brief reversals within Chron C9n may be present at Chilga, thereby providing additional but more indirect support for this correlation. Together these data suggest that this section is probably limited to the duration of Chron C9n (<1 Myr). Vertebrate localities occur primarily through the middle part of the section.

Received 8 May; accepted 30 September 2003; doi:10.1038/nature02102.

- Maglio, V. J. in *Evolution of African Mammals* (eds Maglio, V. J. & Cooke, H. B. S.) 603–619 (Harvard Univ. Press, Cambridge, MA, 1978).
- Murdoch, T. G. *The Chilga Lignite Deposits. US Technical Project in Ethiopia* (Mineral Investigation Memorandum no. 19, Ministry of Mines, Addis Ababa, 1944).
- Yemane, K., Bonnefille, R. & Faure, H. Palaeoclimatic and tectonic implications of Neogene microflora from the Northwestern Ethiopian highlands. *Nature* **318**, 653–656 (1985).
- Yemane, K., Robert, C. & Bonnefille, R. Pollen and clay mineral assemblages of a late Miocene lacustrine sequence from the northwestern Ethiopian highlands. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **60**, 123–141 (1987).
- Yemane, K., Taieb, M. & Faure, H. Limnogeologic studies on an intertrappean continental deposit from the northern Ethiopian Plateau ($37^{\circ}03'$ E, $12^{\circ}25'$ N). *J. Afr. Earth Sci.* **6**, 91–101 (1987).
- Spell, T. L., McDougall, I. & Douleris, A. P. Cerro Toledo rhyolite, Jemez volcanic field, New Mexico: ⁴⁰Ar/³⁹Ar geochronology of eruptions between two caldera-forming events. *Geol. Soc. Am. Bull.* **108**, 1549–1566 (1996).
- Feseha, M. *Sequence Stratigraphy, Petrography, and Geochronology of the Chilga Rift Basin Sediments, Northwest Ethiopia* Thesis, Univ. Texas (2001).
- Cande, S. C. & Kent, D. V. Revised calibration of the geomagnetic polarity timescale for the Late Cretaceous and Cenozoic. *J. Geophys. Res.* **100**, 6093–6095 (1995).
- Hofmann, C. *et al.* Timing of the Ethiopian flood basal event and implications for plume birth and global change. *Nature* **389**, 838–841 (1997).
- Rochette, P. *et al.* Magnetostratigraphy and timing of the Oligocene Ethiopian traps. *Earth Planet. Sci. Lett.* **164**, 497–510 (1998).
- Ukstins, I. A. *et al.* Matching conjugate volcanic rifted margins: ⁴⁰Ar/³⁹Ar chronostratigraphy of pre- and syn-rift bimodal flood volcanism in Ethiopia and Yemen. *Earth Planet. Sci. Lett.* **198**, 289–306 (2002).
- Baker, J., Snee, L. & Menzies, M. A brief Oligocene period of flood volcanism in Yemen: implications for the duration and rate of continental flood volcanism at the AfroArabian triple junction. *Earth Planet. Sci. Lett.* **138**, 39–55 (1996).
- Andrews, C. W. *A Descriptive Catalogue of the Tertiary Vertebrata of the Fayum, Egypt* 1–324 (British Museum of Natural History, London, 1906).
- Tanner, L. G. in *Evolution of African Mammals* (eds Maglio, V. J. & Cooke, H. B. S.) 279–283 (Harvard Univ. Press, Cambridge, Massachusetts, 1978).
- Wight, A. W. R. in *The Geology of Libya* Vol. 1 (eds Salem, M. J. & Busrewil, M. T.) 309–325 (Academic, London, 1980).
- Pickford, M. Première découverte d'une faune mammalienne terrestre paléogène d'Afrique subsaharienne. *Comptes Rendus l'Acad. Sci. Paris Série II* **19**, 1205–1209 (1986).
- Thomas, H. *et al.* in *Fossil Vertebrates of Arabia* (eds Whybrow, P. J. & Hill, A.) 430–442 (Yale Univ. Press, New Haven, 1999).
- Rasmussen, D. T. & Simons, E. L. New Oligocene hyracoids from Egypt. *J. Vert. Paleontol.* **8**, 67–83 (1988).
- Rasmussen, D. T. in *The Evolution of Perissodactyls* (eds Prothero, D. R. & Schoch, R. M.) 57–78 (Oxford Univ. Press, Oxford, 1989).
- Gheerbrant, E., Sudre, J. & Cappetta, H. A Palaeocene proboscidean from Morocco. *Nature* **383**, 68–70 (1996).

- Gheerbrant, E. *et al.* A new large mammal from the Ypresian of Morocco: evidence of surprising diversity of early proboscideans. *Acta Palaeontol. Pol.* **47**, 493–506 (2002).
- Coppens, Y., Maglio, V. J., Madden, C. T. & Beden, M. in *Evolution of African Mammals* (eds Maglio, V. J. & Cooke, H. B. S.) 336–367 (Harvard Univ. Press, Cambridge, 1978).
- Harris, J. M. in *Evolution of African Mammals* (eds Maglio, V. J. & Cooke, H. B. S.) 315–332 (Harvard Univ. Press, Cambridge, 1978).
- Court, N. A new species of *Numidotherium* (Mammalia: Proboscidea) from the Eocene of Libya and the early phylogeny of the Proboscidea. *J. Vert. Paleontol.* **15**, 650–671 (1995).
- Tassy, P. *La Place des Mastodontes Miocènes de l'Ancien Monde dans la Phylogénie des Proboscidea (Mammalia): Hypothèses et Conjectures*. PhD thesis, Mémoires des Sciences de la Terre. Univ. Pierre et Marie Curie, Paris **85-34**, 1–862 (1985).
- Pickford, M. Cainozoic paleontological sites of Western Kenya. *Münchner Geowiss. Abhandlun.* **8**, 1–151 (1986).
- Sanders, W. J. & Miller, E. R. New proboscideans from the early Miocene of Wadi Moghara. *Egypt. J. Vert. Paleontol.* **22**, 388–404 (2002).
- Shoshani, J., Walter, R. C., Libsekal, Y., Abrahama, M. & Berhe, S. in *Sci. Prog. Abstr. 8th Int. Theriological Congr.* 128–129 (Johannesburg, 2001).
- Boschetto, H. B., Brown, F. H. & McDougall, I. Stratigraphy of the Lothidok Range, northern Kenya, and K/Ar ages of its Miocene primates. *J. Hum. Evol.* **22**, 47–71 (1992).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank the Authority for Research and Conservation of Cultural Heritage, the Ministry of Culture and Sports Affairs, Ethiopia, for permission to conduct our ongoing research in the Blue Nile Basin, the Director and staff of the National Museum, Addis Ababa, for their assistance with collections, and the Gondar ARCCCH for logistical support. The project is funded by grants from the National Science Foundation, the National Geographic Society, the L. S. B. Leakey Foundation, and the University Research Institute of UT Austin. W.J.S. was funded in part by a Scott Turner Award from the Department of Geological Sciences of The University of Michigan. A. Asfaw, M. Birara, F. Chanic, A. Kebede and T. Selassie provided field assistance, and M. Crawford and A. Neuenschwander aided with the interpretation of the satellite imagery. B. Miljour drafted Fig. 3. We are grateful to J.-J. Jaeger for comments and insights. We thank the people of the Chilga region for their hospitality and assistance in the fieldwork.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to J.K. (jkappelman@mail.utexas.edu).

Ocean currents mediate evolution in island lizards

Ryan Calsbeek¹ & Thomas B. Smith^{1,2}

¹Center for Tropical Research, Institute of the Environment, ²Organismic Biology, Ecology and Evolution, University of California, Los Angeles, California 90065, USA

Islands are considered to be natural laboratories in which to examine evolution because of the implicit assumption that limited gene flow allows tests of evolutionary processes in isolated replicates¹. Here we show that this well-accepted idea requires re-examination. Island inundation during hurricanes can have devastating effects on lizard populations in the Bahamas^{2,3}. After severe storms, islands may be recolonized by over-water dispersal of lizards from neighbouring islands³. High levels of gene flow may homogenize genes responsible for divergence, and are widely viewed as a constraining force on evolution^{4,5}. Ultimately, the magnitude of gene flow determines the extent to which populations diverge from one another, and whether or not they eventually form new species^{6,7}. We show that patterns of gene flow among island populations of *Anolis* lizards are best explained by prevailing ocean currents, and that over-water dispersal has evolutionary consequences. Across islands, divergence in fitness-related morphology decreases with increasing gene flow⁵. Results suggest that over-water dispersal after hurricanes constrains adaptive diversification in *Anolis* lizards, and that it may have an important but previously undocumented role in this classical example of adaptive radiation.

Anolis lizards in the Caribbean represent one of the best examples of a vertebrate adaptive radiation^{8–10}. Over 140 species of Caribbean anoles have radiated into six morphological classes (termed ‘ecomorphs’)⁸ that are specialized for different habitat types^{11,12}. Ecomorphs that perch on broad-diameter substrates (for example, tree trunks) tend to have longer limbs compared with ecomorphs that perch on narrow-diameter substrates (for example, twigs; see below), and based on locomotor performance, these differences are adaptive¹¹. Phylogenetic evidence has shown that morphological classes evolved independently on different islands in the Greater Antilles¹³ and that islands were colonized by a common ancestor from Cuba, which subsequently radiated into the diversity of species found today.

Recent work indicates that hurricanes have important effects on *Anolis* demography^{2,3}. During September 1999 the category IV hurricane Floyd devastated island lizard populations in the Bahamas with winds exceeding 250 km h⁻¹ and associated tidal surges of over 6 m; populations on 66 study islands were either killed or washed off the islands by the storm³. Schoener *et al.*³ suggested that the initial recovery of these island lizard populations (approximately 2 months after hurricane Floyd) was by progeny hatching from eggs laid before the storm. Although no islands were reported to have received immigrants as a result of hurricane transport, subsequent recolonization of islands over the next 17 months was rapid and indicated over-water dispersal of adult lizards from neighbouring islands³. Here we use microsatellite markers to show that weather-related abiotic factors, including prevailing winds and ocean currents, probably determine the direction of gene flow between islands and consequently the magnitude of

morphological divergence by natural selection. Our study provides genetic evidence that supports the mechanism of island recolonization after hurricanes that was proposed by Schoener *et al.*³, and thus indicates that the joint action of hurricanes and ocean currents may have an important influence on the adaptive radiation of this group.

Variation in microsatellite allele frequencies was used to test the hypothesis that gene flow between insular populations of *Anolis sagrei* occurs by over-water dispersal on ocean currents in the Bahamas. We tested whether differences in the magnitude of gene flow between islands could affect microevolutionary patterns of morphological diversification in this group. The adaptive radiation of Caribbean anoles is believed to be driven by ecologically based natural selection arising from variation in habitat use^{9,10,14}. Selection is thought to operate on traits such as limb length through locomotor performance associated with differences in perch diameter and perch height¹¹. Longer limbs are favoured on broad perching surfaces because long limbs increase maximum sprint speed, allowing lizards to catch prey or escape predators more efficiently. However, shorter limbs are favoured on narrow perches because they enhance agility relative to longer limbs^{11,15}. Numbers of toe-pad lamellae are also correlated with perching height, owing to the importance of clinging ability¹⁶. Interspecific variation in these three characters features prominently in divergence of ecomorphs and consequently in the adaptive radiation of this group^{9,10}. Moreover, intraspecific variation in *A. sagrei* mirrors that of the adaptive radiation of anoles, suggesting that ecologically based natural selection acts similarly at the population and species levels¹⁷.

Prevailing winds and surface ocean currents in the Bahamas have a net westerly flow among islands^{18,19}, except where the Florida

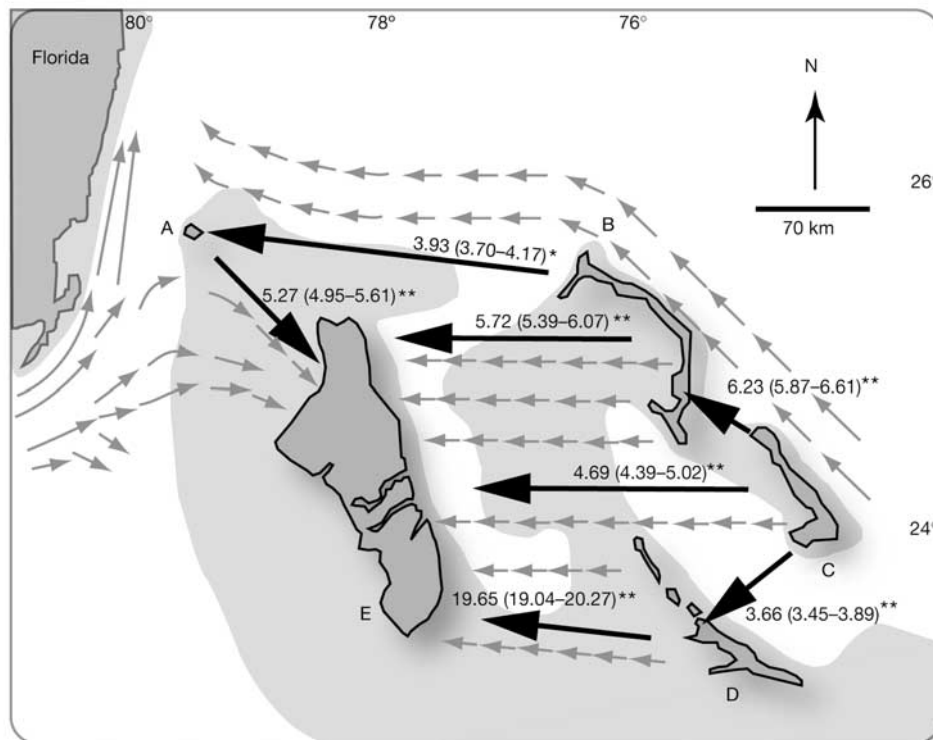


Figure 1 Map of the study islands showing the polarity and magnitude of ocean currents (thin arrows) and gene flow (thick arrows) between adjacent islands. The lightly shaded areas indicate shallow banks surrounding the study islands and Florida. Numbers of migrants per generation (N_m) and 5% and 95% confidence intervals are reported for each pair of adjacent islands in the dominant direction of gene flow as computed in MIGRATE. Confidence intervals indicate N_m values that are significantly greater in the direction shown than in the reverse direction based on maximum likelihood estimates from

MIGRATE. Note that the direction of gene flow is congruent everywhere with the flow of ocean currents. In the case of islands C and D, the juxtaposition of islands relative to ocean currents may account for comparatively small N_m values. Islands are labelled as follows: A, South Bimini; B, Eleuthera; C, Cat Island; D, Great Exuma; E, Andros. Asterisk, non-overlapping 0.5% and 99.5% confidence intervals; double asterisk, non-overlapping 5% and 95% confidence intervals.

current meets the western boundary of the Caribbean current (Fig. 1). Here, the current has a northerly flow around the eastern side of Florida, and an intermittent southward flow from Bimini to Andros islands in the Bahamas²⁰. The deep Florida current cannot flow over the shallow banks of the Bahamas, but flows northward to the Gulf Stream. By contrast, the southward flow from Bimini to Andros islands (Fig. 1) might be driven by northwesterly trade winds associated with the passage of cold fronts during winter months, which could force southeastward surface currents over the shallow banks of the Bahamas. This suggests that propagules (including *Anolis* lizards) introduced to the Atlantic Ocean around the Bahamas should also drift from east to west among most islands, and south from Bimini to Andros.

During June and July of 2002 we visited five islands on the Great Bahama bank and captured approximately 50 *A. sagrei* lizards from each island. We assessed variation at eight microsatellite loci in the laboratory. Maximum likelihood estimates of gene flow direction based on Markov chain Monte-Carlo simulations²¹ indicated that the direction of gene flow among islands is best explained by patterns of ocean currents in the Caribbean, and has occurred predominantly from east to west in the Bahamas (Rayleigh test of directionality $r = 0.76$, $P < 0.008$, Fig. 1; see ref. 22). The average distance among islands in our study is 124 km (Fig. 1). Assuming an average ocean-surface flow velocity of 20 cm s^{-1} (ref. 19), floating or rafting between islands would require approximately 1 week (range 2–15 days). This is a conservative estimate because wind and wave velocities increase substantially during periods of intense storm activity. Moreover, lizard migration could proceed in a stepping-stone fashion, where intermediate islands would reduce exposure time in the open ocean. Laboratory experiments indicate that *A. sagrei* are capable of surviving long-term exposure to sea water (for example, $>24 \text{ h}$ unaided by rafts)²³, and field observations indicate that under some conditions *A. sagrei* willfully enter the ocean and swim away from small, inhospitable islands²⁴. Thus, these results support the assertion of Schoener *et al.*³ that over-water dispersal may be a significant factor in the ecology and evolution of

lizard populations in the Caribbean.

A multivariate estimate of shape (see Methods) based on four morphological characters, including those considered most likely affected by natural selection, revealed that levels of gene flow between islands were significantly correlated with population-level divergence in morphology (Mantel's $r = 1.23$, $P < 0.02$; Fig. 2). One explanation for this correlation is that gene flow has homogenizing effects on morphology and adaptive divergence by natural selection. An alternative explanation, that gene flow and morphological divergence are both correlated with geographical distance between islands, is not supported by our data. There was no relationship between morphological distance and geographical distance between our study populations ($r^2 = 0.019$, $P = 0.40$). Gene flow between islands could therefore have important implications for the adaptive radiation of this group by constraining the ability of natural selection to drive morphological divergence. Moreover, although some of the variation in fitness-related characters may arise by plastic changes in response to the environment (for example, differences in perch diameter)²⁵, plasticity alone cannot explain the significantly negative relationship between divergence and gene flow. Gene flow would not be expected to have homogenizing effects on divergence among populations in a trait that was completely plastic²⁶. The homogenizing effects of gene flow on morphological divergence reported here indicate that some of the phenotypic variation probably arises owing to differences among individuals at genetic loci controlling limb length.

Alternative explanations for the directional patterns of gene flow reported in this study include human transportation of lizards among islands, and historical contingencies such as land bridges that connected islands of the Great Bahama bank during the Last Glacial Maximum (approximately 8,000–10,000 yr before present)²⁷ or island colonization from Cuban ancestors. These are unlikely explanations. First, neither human-mediated dispersal of anoles nor historical connections among islands would be expected to produce dominantly unidirectional patterns of gene flow. In contrast, we find directionality of gene flow that is congruent in all cases with the prevailing direction of ocean currents, including the exceptional case in which currents adjacent to Florida drive gene flow south from Bimini to Andros islands (Fig. 1). Second, *Anolis* populations in the Bahamas are derived from Cuban anoles located to the southwest of the Bahamas. Thus colonization from Cuba should have produced a directionality of gene flow exactly opposite to that reported here (that is, west to east rather than east to west). We therefore conclude that transport on ocean currents remains the most parsimonious explanation for the highly directional patterns of gene flow among islands.

Studies on islands have revealed many of the fundamental mechanisms of evolution, particularly the paramount influence of geographical isolation to diversification¹. Here, we add an important caveat to these studies, showing that prevailing ocean currents may influence gene flow and adaptive divergence in a terrestrial vertebrate. The adaptive radiation of anoles in the Caribbean is thought to have arisen by ecologically based natural selection related to habitat use²⁸. However, the level of gene flow between populations will impose an upper limit on the ability of natural selection to drive adaptive divergence⁵. We have provided evidence that weather-related abiotic phenomena might have important effects on the evolution and adaptive radiation of lizard populations. □

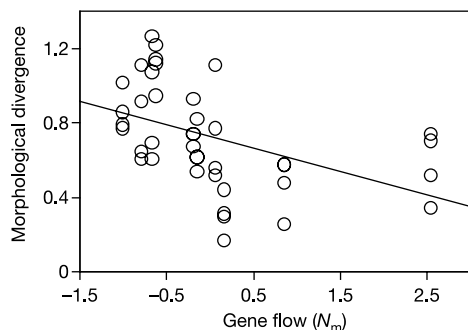


Figure 2 Standardized morphological divergence (mean = 0, unit variance) in four morphological characters, as a function of the standardized number of migrants (N_m) between islands. Morphological divergence decreased with increasing gene flow, showing that migration between islands slows the rate of adaptive divergence due to natural selection. The significant negative slope indicates that the difference in trait divergence has a genetic basis and is not simply due to environmental factors²⁵. We calculated trait divergence as the absolute value of the difference in euclidean distance between the following standardized characters: hindlimb length, forelimb length, toe-pad lamella width and snout-vent length. We calculated morphological divergence within and between habitat types (broad- and narrow-perch diameters) and among island populations. Each pair of islands generated four morphological comparisons, two each within and between habitat types. Thus, each population was used in multiple morphological comparisons and we used pairwise matrix correlations (Mantel's test) to account for non-independence among population comparisons. We present the least-squares regression line for illustrative purposes only.

Methods

Field

We captured 45–52 lizards from each island (mean of 51) by hand or using a silk noose. Lizards were captured at three to five locations along linear transects on each island (end points separated by 20 km except on South Bimini where small island size constrained transect length to 6 km). Animals were sexed, weighed and measured. Hindlimb and forelimb lengths were measured (to nearest millimetre) from the point of insertion at the abdomen to the distal tip of the femoral-tibial and humero-radio-ulnar joints. Toe-pad lamella width was measured with dial calipers (average of three measurements) at the

widest point of the distal toe pad on the fourth toe of the hind foot. We collected a 2-mm piece of tissue from the distal tip of the tail for molecular analyses. All lizards were marked with a small spot of white paint to prevent re-capture and were immediately released to their original point of capture.

Morphology

We regressed natural-log-transformed limb lengths and toe-pad width on snout-vent length (following procedure of ref. 29) and considered the residuals about the regression line to represent size-independent morphology. Lizards captured on broad-diameter (>15 cm) perches had significantly longer hind- and forelimbs (mean $\pm$ standard error) relative to body size compared with lizards captured on narrow-diameter (<15 cm) perches (residual hindlimb_{broad} = 0.028 ± 0.10 versus residual hindlimb_{narrow} = -0.009 ± 0.007 , $P < 0.002$; residual forelimb_{broad} = 0.029 ± 0.11 versus residual forelimb_{narrow} = -0.015 ± 0.10 , $P < 0.006$). This difference was also significant when we controlled for snout-vent length in an analysis of covariance¹⁷.

To reduce the dimensionality of our data, we combined the three fitness-related morphological characters and snout-vent length (that is, body size) into a unique, multivariate estimate of shape by calculating the euclidean distance between the four traits using SYSTAT version 5.2.1. Morphological divergence between islands was calculated using mean values for morphological traits from each island population. Before calculating means, data for all variables were standardized to mean zero with unit variance. The significance of the correlation between morphological divergence and gene flow was calculated using pairwise matrix correlations (Mantel's test) in the statistical package 'R' version 4.0. We computed the test statistic using two 10×10 matrices representing populations from two habitat types on each of five islands. Significance level was computed with 9,999 permutations (as recommended by ref. 30).

Genetics

Genomic DNA was extracted from tail tissue by overnight incubation at 55 °C in 500 μ l 5% Chelex (Biourad) and 2 μ l proteinase K solution (at 20 mg ml⁻¹) followed by centrifugation and 1:10 dilution of the extract. We amplified eight loci from the genomic template via polymerase chain reaction, and length polymorphism among individuals (40–44 lizards per island) was assessed with fluorescent labelling of one of the primers on an automated DNA sequencer (ABI 3700).

We estimated the magnitude and polarity of gene flow using the maximum likelihood approach implemented by MIGRATE-1.6.9. MIGRATE extends a coalescence approach to include migration rates among populations, assuming a per-locus mutation rate μ , and constant population sizes. This approach has been used in numerous studies and is considered to estimate gene flow more accurately than other Fst methods²¹. Results reported here use 50 short-chain searches and three long-chain searches over eight microsatellite loci using a brownian motion approximation to the ladder model of evolution. This search strategy resulted in the best likelihood scores, although other runs using fewer short-chain searches produce qualitatively similar results (N_m values from all runs were highly correlated and we report mean values with confidence intervals). N_m values represent the per-generation migration rate between the study islands depicted in Fig. 1 and are reported for the direction of greater magnitude between each pair of adjacent islands.

Received 25 August; accepted 9 October 2003; doi:10.1038/nature02143.

- Mayr, E. *Animal Species and Evolution* (Belknap, Cambridge, Massachusetts, 1963).
- Spiller, D. A., Losos, J. B. & Schoener, T. W. Impact of a catastrophic hurricane on island populations. *Science* **281**, 695–697 (1998).
- Schoener, T. W., Spiller, D. & Losos, J. B. Natural restoration of the species-area relation for a lizard after a hurricane. *Science* **294**, 1525–1528 (2000).
- Endler, J. A. *Geographic Variation Speciation and Clines* (Princeton Univ. Press, Princeton, New Jersey, 1977).
- Slatkin, M. Gene flow and the geographic structure of natural-populations. *Science* **236**, 787–792 (1987).
- Rice, W. R. & Hostert, E. E. Laboratory experiments on speciation: What have we learned in 40 years? *Evolution* **47**, 1637–1653 (1993).
- Barton, N. H. & Hewitt, G. M. Adaptation, speciation, and hybrid zones. *Nature* **341**, 497–503 (1989).
- Williams, E. E. in *Lizard Ecology: Studies of a Model Organism* (eds Huey, R. B., Pianka, E. R. & Schoener, T. W.) 326–370 (Harvard Univ. Press, Cambridge, Massachusetts, 1983).
- Losos, J. B., Jackman, T. R., Larson, A., DeQueiroz, K. & Rodriguez-Shettino, L. Contingency and determinism in replicated adaptive radiations of island lizards. *Science* **279**, 2115–2118 (1998).
- Schluter, D. in *The Ecology of Adaptive Radiation* (eds May, R. M. & Harvey, P. H.) (Oxford Univ. Press, New York, 2000).
- Losos, J. B. The evolution of form and function: Morphology and locomotor performance in West Indian *Anolis* lizards. *Evolution* **44**, 1189–1203 (1990).
- Irschick, D. J. & Losos, J. B. A comparative analysis of the ecological significance of maximal locomotor performance in Caribbean *Anolis* lizards. *Evolution* **52**, 219–226 (1998).
- Jackman, T. R., Larson, A., De Queiroz, K. & Losos, J. B. Phylogenetic relationships and tempo of early diversification in *Anolis* lizards. *Syst. Biol.* **48**, 254–285 (1999).
- Irschick, D. J., Vitt, L. J., Zani, P. A. & Losos, J. B. A comparison of evolutionary radiations in mainland and Caribbean *Anolis* lizards. *Ecology* **78**, 2191–2203 (1997).
- Losos, J. B. & Sinervo, B. The effect of morphology and perch diameter on sprint performance of *Anolis* lizards. *J. Exp. Biol.* **145**, 23–30 (1989).
- Irschick, D. J. *et al.* A comparative analysis of clinging ability among pad-bearing lizards. *Biol. J. Linn. Soc.* **59**, 21–35 (1996).
- Calsbeek, R. & Smith, T. B. Intraspecific variation in *Anolis sagrei* mirrors the adaptive radiation of Caribbean anoles. *Am. Nat.* (submitted).
- Lee, T. N., Johns, W. E., Zantopp, R. J. & Fillenbaum, E. R. Moored observations of western boundary current variability and thermohaline circulation at 26.5 degrees North in the subtropical North

Atlantic. *J. Phys. Oceanogr.* **26**, 962–983 (1996).

- Kinder, T. G., Heburn, G. W. & Green, A. W. Some aspects of the Caribbean circulation. *Mar. Geol.* **68**, 25–52 (1985).
- Georges, T. M., Harlan, J. A. & Lematta, R. A. Large-scale mapping of ocean-surface currents with dual over-the-horizon radars. *Nature* **379**, 434–436 (1996).
- Beerli, P. & Felsenstein, J. Maximum likelihood estimation of a migration matrix and effective population sizes in n subpopulations by using a coalescent approach. *Proc. Natl Acad. Sci. USA* **98**, 4563–4568 (2001).
- Batchelet, E. *Circular Statistics in Biology* (Academic, Toronto, 1981).
- Schoener, A. & Schoener, T. W. Experiments on dispersal: Short-term flotation of insular anoles, with a review of similar abilities in other terrestrial animals. *Oecologia (Berl.)* **63**, 289–294 (1984).
- Schoener, T. W. & Schoener, A. in *Advances in Herpetology and Evolutionary Biology* (eds Rhodin, A. G. & Miyata, K.) 491–498 (Museum of Comparative Zoology, Harvard University, 1983).
- Losos, J. B. *et al.* Evolutionary implications of phenotypic plasticity in the hindlimb of the lizard *Anolis sagrei*. *Evolution* **54**, 301–305 (2000).
- Smith, T. B., Wayne, R. K., Girman, D. J. & Bruford, M. W. A role for ecotones in generating rainforest biodiversity. *Science* **276**, 1855–1857 (1997).
- Lighty, R. G., Macintyre, I. G. & Stuckenrath, R. Holocene reef growth on the edge of the Florida Shelf. *Nature* **278**, 281–282 (1979).
- Ogden, R. & Thorpe, R. S. Molecular evidence for ecological speciation in tropical habitats. *Proc. Natl Acad. Sci. USA* **99**, 13612–13615 (2002).
- Losos, J. B. Ecomorphology, performance capability, and scaling of West Indian *Anolis* lizards: An evolutionary analysis. *Ecol. Monogr.* **60**, 369–388 (1990).
- Casgrain, P., Legendre, P. & Valdour, A. *The R Package for Multivariate and Spatial Analysis, Version 4.0 d5-User's Manual* (<http://www.fas.umontreal.ca/BIOL/legendre/>) (Department de sciences biologiques, Université de Montreal, Canada, 2001).

Acknowledgements We thank P. Beerli and N. Manoukis for technical support using Migrate. T. Georges, J. Harlan and T. Lee provided helpful discussions concerning the action of ocean currents. D. Irwin, J. Levine, J. Losos, B. Mila and B. Sinervo provided helpful comments that improved the manuscript. We thank the Commonwealth of the Bahamas for permission to conduct research, and NSF and UCLA for financial support.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to R.C. (calsbeek@ucla.edu).

Delta-promoted filopodia mediate long-range lateral inhibition in *Drosophila*

Cyrille de Jossineau, Jonathan Soulé*, Marianne Martin*, Christelle Anguille, Philippe Montcourrier & Daniel Alexandre

Laboratoire Dynamique Moléculaire des Interactions Membranaires, UMR 5539, Université Montpellier II C.C. 107, Place Eugène Bataillon, 34095 Montpellier cedex 5, France

*These authors contributed equally to this work

Drosophila thoracic mechanosensory bristles originate from cells that are singled out from 'proneural' groups of competent epithelial cells. Neural competence is restricted to individual sensory organ precursors (SOPs) by Delta/Notch-mediated 'lateral inhibition', whereas other cells in the proneural field adopt an epidermal fate. The precursors of the large macrochaetes differentiate separately from individual proneural clusters that comprise about 20–30 cells or as heterochronic pairs from groups of more than 100 cells¹, whereas the precursors of the small regularly spaced microchaetes emerge from even larger proneural fields². This indicates that lateral inhibition might act over several cell diameters; it was difficult to reconcile with the fact that the inhibitory ligand Delta is membrane-bound until the observation that SOPs frequently extend thin processes^{3,4} offered an attractive hypothesis. Here we show that the extension of these planar filopodia—a common attribute of wing imaginal disc cells—is promoted by Delta and that their experimental suppression reduces Notch signalling in distant cells and

increases bristle density in large proneural groups, showing that these membrane specializations mediate long-range lateral inhibition.

Long-range action of lateral inhibition has been postulated since the pioneering work of Wigglesworth⁵ and has been supported by indirect evidence^{4,6}. We assessed the planar range of Delta–Notch signalling by probing for expression of the Enhancer of split m8 (E(spl)m8) gene, a direct downstream transcriptional target of the Notch pathway^{7,8}. When Delta is overexpressed in macrochaete SOPs, the number of E(spl)m8-expressing neighbouring cells increases by 50–80% and this affects cells located about three to five cell diameters away from the precursor (compare Fig. 3c, f, i with Fig. 3b, e, h, respectively). As shown by the use of a Delta::LacZ transgene, Delta transcription is not triggered in these cells but only in the Delta-overexpressing SOP itself (data not shown). Because the only other known Notch agonist, Serrate, has no effect on lateral inhibition⁹, this rules out a relay mechanism by Notch ligands and suggests that lateral inhibition operates at a distance in a direct way by transport. The finding of a soluble cleavage product of Delta consisting of its extracellular part suggested that transport might proceed by diffusion¹⁰, but this ligand fragment was subsequently found to be unable to activate the Notch pathway¹¹. To study Delta transport, we constructed a tagged version of Delta in which green fluorescent protein (GFP) was inserted at the distal end of the extracellular domain, just after the conceptual signal peptide. As judged by mutation complementation and overexpression analysis, this construct behaves as a gain of function (data not shown), indicating that its signalling activity is maintained. When expressed in macrochaete SOPs, this fusion protein localized to vesicular structures and at the plasma membrane, as the unmodified Delta protein does¹². Most notably, the previously described SOP filopodia^{3,4} were also stained (Fig. 1d), and when the SOP outline was highlighted with a membrane-addressed CD8–GFP fusion protein, Delta immunostaining labelled the majority of these processes, especially on the apical and basal aspects (Fig. 1a–c and ref. 4), showing that these structures mediate the planar transport of Delta. A similar labelling of cellular extensions was observed when CD8–GFP was expressed in randomly selected individual cells by use of the ‘flip-on’ method¹³ (Fig. 1e). About 30% of the columnar cells in the late-third-instar wing imaginal disc proper extend planar processes of one cell diameter or more in length. These filopodia-displaying cells are seemingly randomly distributed and we could not link their occurrence to any specific location. In fact, 1–20- μ m-long filopodia seem to be a common feature of imaginal disc cells¹⁴ rather than a specific attribute of SOPs, but we further studied these cellular processes exclusively in SOPs, taking advantage of the invariant position of these individually identifiable cells¹.

When CD8–GFP is expressed in SOPs, imaging the intrinsic fluorescence of GFP in live imaginal discs reveals only a few planar extensions (data not shown), whereas when stained with an anti-GFP antibody, their number and length seem significantly larger, with filopodia being frequently branched and varying in diameter along their proximodistal axis (Fig. 1f). They can be observed at various levels along the apico-basal axis, but most are seen on the basal side, where they resemble the epidermal feet described in various insects¹⁵, and on the apical surface, where they extend from the broadened cortex of the cell and run above the zonula adherens between or over neighbouring cells. Microchaete precursors similarly display filopodial extensions (Fig. 2d and ref. 4), which are mainly oriented along the antero-posterior axis of the pupae. We propose that *in vivo* microscopy allows imaging of only the thickest and most strongly labelled proximal part of cellular processes, whereas immunostaining reveals a larger fraction of the filopodia web, most probably as a result of an improved signal to noise ratio. However, it should be stressed that the thinnest processes are possibly beyond the resolving power of our imaging setup, even after immunostaining. We noticed incidentally that even slight

compression of the tissue under the coverslip resulted in a significantly larger filopodia web (data not shown; see Methods), as though the tension applied to the mechanically constrained epithelial cell was relaxed by the growth of processes in the less constrained environment of the basal and apical surfaces. Whether this phenomenon is an artefact or reflects a morphogenetic role of actual tissue tensions remains an open but interesting issue. With these potential caveats in mind, we studied the dynamics of SOP processes. All SOPs exhibit filopodia and their extension follows a reproducible sequence (Fig. 1g–i). Only a few short processes are visible on young SOPs (Fig. 1g), although this should be interpreted with caution because the low and recent activity of the neuralized promoter might impede their labelling. Filopodia soon reach a fixed degree of extension (Fig. 1h), and we noticed that SOPs emerging in the vicinity of an already determined precursor from the same proneural group always developed out of reach of its visible filopodia web (Fig. 1i) even if contacts were observed in later stages. In these heterochronic pairs¹, the extension of processes from the young SOP was concomitant with some retraction of those of the older precursor (Fig. 1h, i). Time-lapse imaging showed that the observable proximal regions of filopodia are stable over periods of about half an hour because no obvious retraction, elongation, branching or lateral movements were observed in our movies (data not shown). As judged by immunostaining at the puparium formation stage, the overall pattern of the filopodia web and the number of potentially contacted neighbouring cells are stable for a given SOP (Fig. 2c). Filopodia were still apparent after the divisions

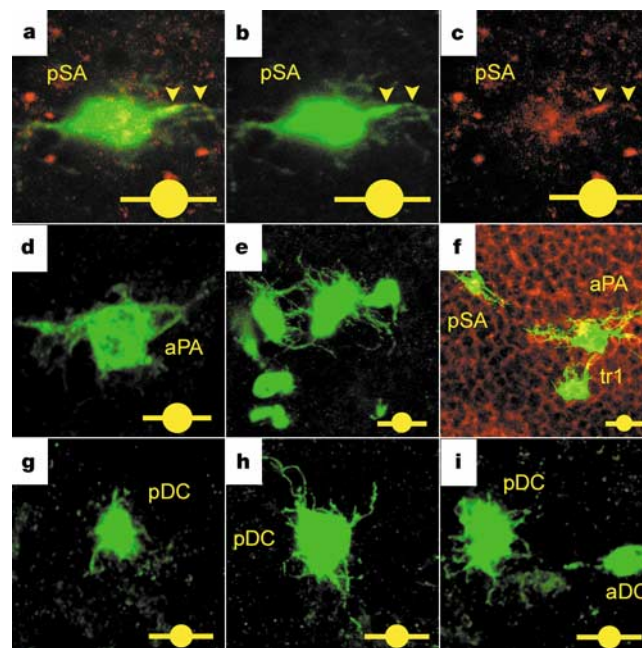
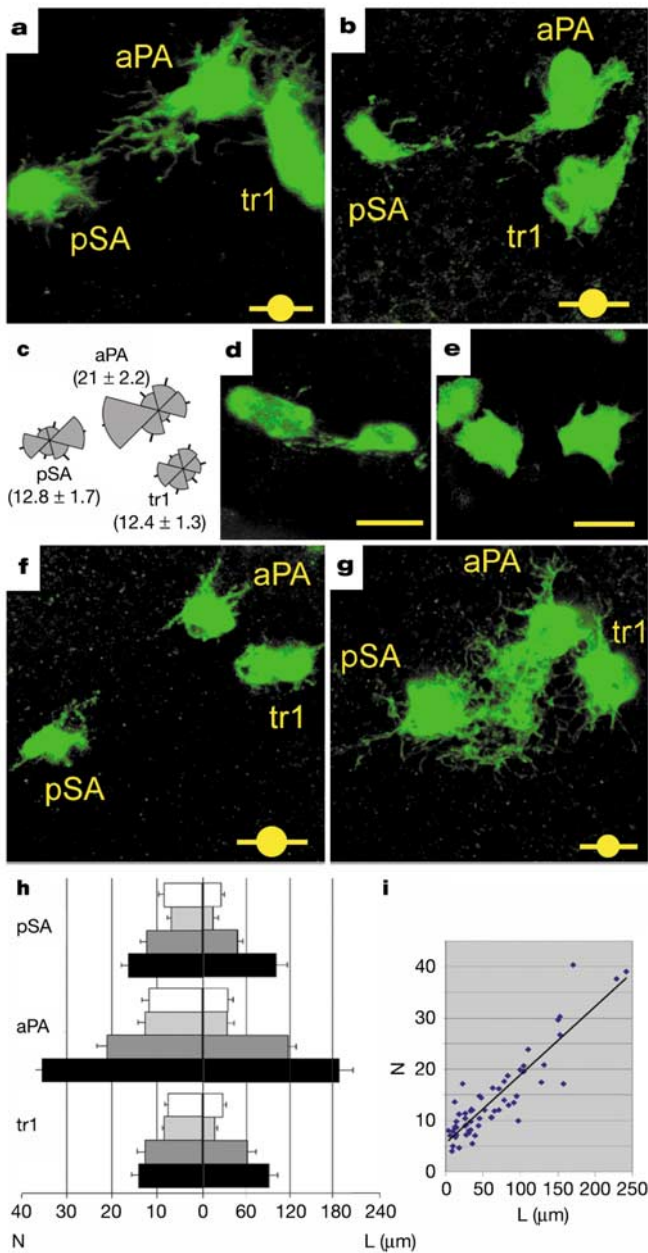
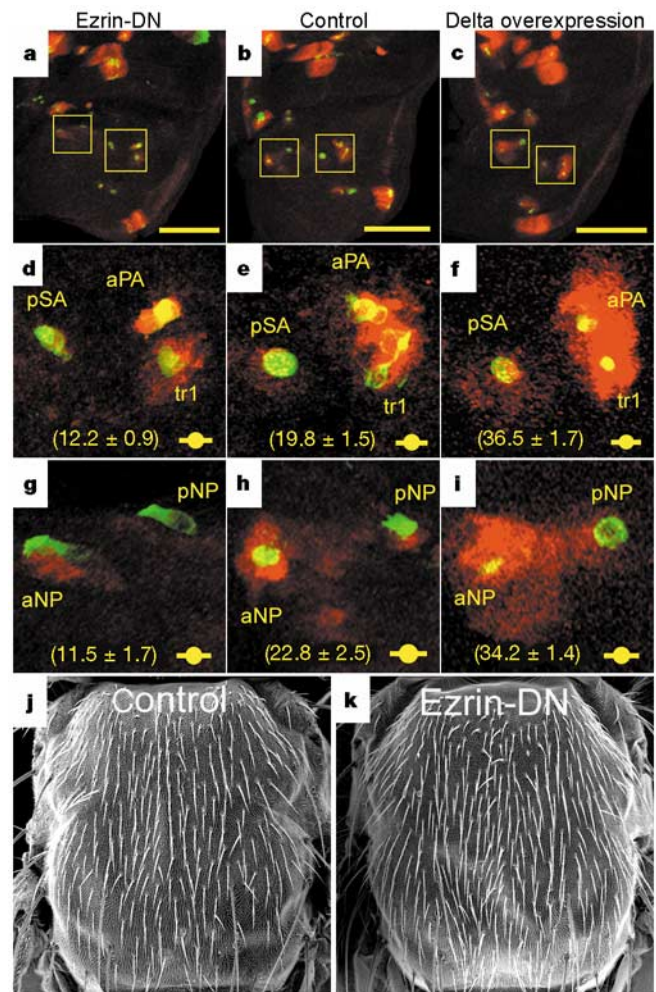


Figure 1 Confocal imaging of anti-GFP immunofluorescent labelling of wing disc cells expressing CD8–GFP (**a–c**, **e–i**) or Delta–GFP (**d**) at the puparium formation stage unless otherwise stated. **a–c**, Basal aspect of a SOP labelled with anti-GFP (green in **a** and **b**) and anti-Delta (red in **a** and **c**). A Delta-stained process is arrowed. **d**, SOP expressing Delta–GFP. **e**, Small random clones of cells expressing CD8–GFP in the notum. **f**, Montage of the maximum projection of three macrochaete precursors (green) superimposed in a single focal plane where cell outlines are labelled with phalloidin–TRITC (red). The terminal broadening of SOPs makes them look larger than the surrounding cells. **g–i**, Dorsocentral precursors 5 h before puparium formation (**g**), 2 h before puparium formation, just before the emergence of the aDC SOP (**h**) and at puparium formation (**i**). Scale bars, 10 μ m. Filled circles show the average diameter of epithelial cells, which varies slightly with genotype. aPA, anterior postalar; tr1, trichoid sensilla 1; pSA, posterior supra-alar; aDC and pDC, anterior and posterior dorsocentral, respectively.



of the precursor, although we could not resolve whether this affected all daughter cells. Finally, we found that SOP processes were significantly reduced in flies heterozygous for a Delta loss-of-function mutation (Fig. 2h; compare with Fig. 2a, f). Conversely, overexpression of Delta in SOPs led to an increase in the extent of their filopodia web (Fig. 2h), resulting in extreme cases in neighbouring cells being covered by a dense netting of intermingled processes (Fig. 2g). This is reminiscent of the situation in cultured cells, where Delta similarly promotes the extension of cellular processes¹⁶. On the one hand this can be viewed as Delta promoting its own means of transport, but on the other hand this shares a key property with mere diffusion in that range varies with concentration at the source.

If these filopodia are involved in mediating long-range Delta



of the precursor, although we could not resolve whether this affected all daughter cells. Finally, we found that SOP processes were significantly reduced in flies heterozygous for a Delta loss-of-function mutation (Fig. 2h; compare with Fig. 2a, f). Conversely, overexpression of Delta in SOPs led to an increase in the extent of their filopodia web (Fig. 2h), resulting in extreme cases in neighbouring cells being covered by a dense netting of intermingled processes (Fig. 2g). This is reminiscent of the situation in cultured cells, where Delta similarly promotes the extension of cellular processes¹⁶. On the one hand this can be viewed as Delta promoting its own means of transport, but on the other hand this shares a key property with mere diffusion in that range varies with concentration at the source.

If these filopodia are involved in mediating long-range Delta

signalling, their experimental suppression would be expected to result in a reduced range of lateral inhibition, leading to an excess of bristles. SOP extensions were suppressed by treatment of imaginal discs with cytochalasin D but were unaffected by nocodazole (data not shown), suggesting that these structures are actin-based. We thus sought to disturb filopodia formation by upsetting interactions between the actin cytoskeleton and the plasma membrane. We tried to interfere with several known regulators of cell architecture such as members of the small GTPases family, but this resulted in pleiotropic phenotypes as the SOPs lost their typical columnar shape in addition to retracting their filopodia (data not shown). In contrast, we found that a dominant-negative form of ERM protein (for ezrin, radixin, moesin) specifically affected the formation of filopodia when expressed in SOPs. ERM proteins are plasma membrane-associated proteins that regulate membrane-microfilament interactions (reviewed in ref. 17). We previously showed that the carboxy-terminal domain of human ezrin induced filopodial extension when expressed in insect cells in culture and that this could be antagonized by co-expressing the amino-terminal domain^{18,19}. We reasoned that this ezrin dominant-negative N-terminal domain (ezrin-DN) might similarly suppress the naturally occurring filopodia of fly SOPs. In fact, cellular processes were greatly reduced whereas epithelial shape was maintained when this construct was expressed in SOPs (Fig. 2h; compare Fig. 2a, d with Fig. 2b, e, respectively). Staining with an anti-Delta antibody showed that the expression level of the ligand in the SOP was unaffected (data not shown). The number of Delta-positive vesicles per cell was indistinguishable from that in controls (7.8 ± 0.919 versus 8.2 ± 1.155 , respectively; $n = 10$) as were plasma membrane levels, indicating that Delta endocytosis was not impaired by ezrin-DN expression. Sensory organs originate from SOPs by a stereotyped series of asymmetric divisions in which Notch signalling between daughter cells specify their distinct fates. In this process, Delta and Serrate are redundant in activating the Notch pathway and the surrounding epithelial cells cannot compensate for their failure⁹. As judged by staining for ELAV and GFP at 40 h after puparium formation, the fixed sensory organ lineage proceeded normally from SOPs expressing ezrin-DN, and the frequency of fate transformations was similar to that in controls ($2.47 \pm 0.29\%$ and $2.61 \pm 0.57\%$, respectively; $n = 500$), showing that the construct allowed cell division and Notch signalling between immediate neighbouring cells. Although pleiotropy can never be formally excluded, we concluded that ezrin-DN reduced the span of the SOP filopodial web without detectably affecting the other known essential aspects of SOP biology. Ezrin-DN reduced the cumulative length of filopodia per cell by 70–75% and the number of potentially contacted cells by 39% on average (Fig. 2h, i). Use of the E(spl)m8::LacZ construct showed that expressing ezrin-DN in SOPs resulted in a 50% decrease in the number of β -Gal-positive cells; the remaining E(spl)m8-expressing cells were confined to the vicinity of SOPs (compare Fig. 3a, d, g with Fig. 3b, e, h, respectively). This reduced range of Notch signalling was associated with a 20% increase in microchaete density (Table 1), indicating that filopodia might be involved in establishing an inhibitory field around microchaete SOPs during their successive waves of determination². This figure increased above 40% when ezrin-DN was expressed earlier and in the whole notum (Table 1; compare Fig. 3k with Fig. 3j), showing

Table 1 Neurogenic phenotypes induced by ezrin-DN: microchaetes

Expression	Genotype	Microchaete density (%)	<i>n</i>
None	UAS::ezrin(1–280)GFP/+; +	2.85 ± 0.22	5
SOP	UAS::ezrin(1–280)GFP/+; Neu::Gal4/+	3.43 ± 0.28	5
Notum	UAS::ezrin(1–280)GFP/+; Klu::Gal4/+	4.06 ± 0.32	5

The number of flies observed is given as *n*. Expression was induced between puparium formation and 14 h after puparium formation (25 °C).

Table 2 Neurogenic phenotypes induced by ezrin-DN: macrochaetes

Expression	Genotype	Frequency of ectopic dorsocentral or scutellar macrochaetes (%)	<i>n</i>
None	+; Neu::Gal4/+	1.1	92
SOP	ezrin(1–280)GFP/+; Neu::Gal4/+	4.4	68
None	Sca::Gal4/+; +	3.3	60
Proneural group	Sca::Gal4/ezrin(1–280)GFP; +	33.1*	130

Asterisk denotes significant difference ($P < 0.05$); *n* represents observed heminota.

that cellular processes of non-SOP cells are also required for the proper spacing of microchaetes. In contrast, macrochaetes were not significantly affected when the construct was expressed in SOPs only, but a third bristle appeared between heterochronic pairs in 30% of the flies in which filopodia were suppressed earlier and in the whole proneural cluster (Table 2). Taken together, these results indicate that cellular processes mediate long-range lateral inhibition between successive waves of precursor determination. However, they do not demonstrate that these structures are involved in the specification of single SOPs from small proneural groups.

According to the currently favoured model of SOP determination, the biochemical network that links proneural gene self-activation to Delta-mediated lateral inhibition²⁰ is topologically analogous to the reaction-diffusion chemical network envisaged by Alan Turing²¹, but for such a network to have long-range pattern-generating properties some sort of long-range inhibition is required. In this perspective, our finding that long-range lateral inhibition is mediated by Delta-promoted filopodia suggests that complex spatial patterns of sensory bristles might emerge from the dynamic behaviour of the lateral inhibition genetic circuit. □

Methods

Flies

The following transgenic lines were used: hs-Flp; tub:FRT/stop/FRT:Gal4 (ref. 13), Neu::Gal4 (ref. 22), UAS::CD8-GFP (ref. 23), klu::Gal4 (ref. 24), E(spl)m8::LacZ (ref. 8), UAS::DI (ref. 25), DI::LacZ (ref. 26), Sca::Gal4 (ref. 24). The thermosensitivity of the Gal4/UAS system was used to control timed expression of the transgenes. Depending on genotypes, expression was kept low at 16–18 °C during early development and increased to 25–30 °C 10–24 h before staining or imaging, unless otherwise stated. The heat-sensitive DI^{RF} mutation²⁷ was maintained at 18 °C and shifted to 30 °C 21 h before staining.

Generation of transgenic flies

Plasmid pEGFP-ezrin²⁸ was used to produce an intermediate construct allowing expression of the first 280 amino acids of ezrin fused to GFP. Briefly, pEGFP-ezrin was digested by AgeI and then partly digested by BamHI. After filling in with Klenow enzyme and blunt-ended ligation with T4 DNA ligase, plasmid pEGFP-ezrin(1–280) was obtained. From this plasmid, fragment EcoRI–NotI was purified and cloned in the corresponding restriction sites of the pUAST vector, allowing GAL4-induced expression of a dominant-negative ezrin(1–280)–GFP fusion protein. A GFP-Delta fusion protein in which GFP is located in the extracellular domain of Delta was constructed by inserting GFP after codon 31 in the Delta cDNA (gift from M. Muskavitch), just after the conceptual signal peptide. For this purpose, GFP S65T was amplified with the oligonucleotides 5'-TGAGTAAAGGAGAAGAAC-3' and 5'-TTTGTATAGTTCATCATGC-3' and cloned into the ScaI site of the Delta cDNA. The resulting construct was transferred as an EcoRI fragment into the corresponding site of the pUAST vector and checked by sequencing.

Immunohistochemistry

We used the following antibodies and dilutions: mouse anti-Delta (directed against the extracellular domain¹⁰; DSHB), 1:600; rabbit anti-GFP (Clontech), 1:800; mouse anti- β -Gal (Promega), 1:600; mouse anti-ELAV (DSHB), 1:250; fluorescein isothiocyanate-labelled goat anti-rabbit IgG (Sigma), 1:800; Texas Red-labelled anti-mouse IgG (Sigma), 1:600. Imaginal discs and dissected pupae were fixed in 2% paraformaldehyde in PBS for 30 min and stained by standard methods. Phalloidin-tetramethylrhodamine β -isothiocyanate (TRITC) (Sigma) was added to the secondary antibody incubation at $2 \mu\text{g ml}^{-1}$.

Confocal microscopy

Live imaginal discs were dissected and mounted, peripodial side up, in modified TC100 culture medium²⁹. Preparations were covered with a no. 1.5 coverslip supported by spots of nail varnish and sealed with a ring of silicon grease enclosing as much air as possible. Imaging was performed with a Leica TCS4D confocal microscope through a 63 $\times$ /1.20 N.A. water-immersion lens with a carefully adjusted correction collar. All preparations were first checked for the absence of tissue compression by X–Z scanning in reflection mode,

because we noticed that the slightest flattening of imaginal discs during mounting or observation resulted in the outgrowth of longer cellular processes. Further compression resulted in cells lying down towards the outside of the disc and stretching tremendously while maintaining basal and apical adhesions before the basal lamina eventually broke. The long and straight processes observed in these conditions and referred to as cytonemes by others³⁰ characteristically point towards the disc centre and are not seen when the tissue is rigidified by aldehyde fixation. It is therefore our belief that these structures are 'couched-and-stretched' cells. Stained preparations were mounted in Citifluor (Agar Scientific) and imaged in the same way with a 25×/0.75, 40×/1.0 or 63×/1.4 oil-immersion lens. Sequential scanning was used for all double labelling and zooming factors, and Z steps were chosen to optimize pixel size so as to match Rayleigh resolution while avoiding aliasing. Post-acquisition median filtering and contrast adjustment of image stacks and movies were performed with Imaris software (Bitplane).

Quantitative image analysis

Photoshop (Adobe) was used to draw one-pixel-wide lines over processes, and their cumulative lengths per cell and per 60° sector were calculated by counting pixels in Image Tool (University of Texas). The number of cells potentially contacted by the filopodia web of a given SOP was counted in phalloidin-stained preparations or, because disc cells typically have five to seven neighbours, estimated by superimposing images with a hexagonal lattice five times at random per cell and in five discs. Because cell diameter varied with genotype, grid cell size was matched to the mean cell diameter, as measured in optical sections. Cells expressing the E(spl)m8-driven LacZ reporter were counted as β-Gal-positive when the fluorescence intensity in the unfiltered image was 20% or more above background. These experiments, in which discs of various genotypes were processed in parallel, were repeated five times. Significance was assessed by a Fisher exact test with the Statview software (Abacus Concepts) and results were considered significant at $P < 0.05$.

Microchaete density and macrochaete counts

Adult flies were killed at -20°C and gold-palladium sputtered at room temperature. Thoraxes were imaged with a Hitachi S4000 scanning electron microscope at 15 kV with a working distance of 25 mm. Metamorph software (Universal Imaging) was used to count microchaetes and epidermal cells in the rectangle defined by the dorsocentral macrochaetes. Microchaete density was expressed as a percentage of the total cells and bristles. Significance was assessed by a t -test for independent samples with Statistica software (Statsoft). Macrochaetes were counted on live flies, and significance was assessed by a Fisher exact test with Genepop software (<http://wbiomed.curtin.edu.au/genepop/>). For both tests, results were considered significant at $P < 0.05$.

Received 10 September; accepted 28 October 2003; doi:10.1038/nature02157.

- Huang, F., Dambly-Chaudière, C. & Ghysen, A. The emergence of sense organs in the wing disc of *Drosophila*. *Development* **111**, 1087–1095 (1991).
- Usui, K. & Kimura, K. Sequential emergence of the evenly spaced microchaetes on the notum of *Drosophila*. *Roux's Arch. Dev. Biol.* **203**, 151–158 (1993).
- Lai, E. C. & Rubin, G. M. Neuralized functions cell-autonomously to regulate a subset of notch-dependent processes during adult *Drosophila* development. *Dev. Biol.* **231**, 217–233 (2001).
- Renaud, O. & Simpson, P. Scabrous modifies epithelial cell adhesion and extends the range of lateral signalling during development of the spaced bristle pattern in *Drosophila*. *Dev. Biol.* **240**, 361–376 (2001).
- Wigglesworth, V. B. Local and general factors in the development of 'pattern' in *Rhodnius prolixus* Hemiptera. *J. Exp. Biol.* **17**, 180–200 (1940).
- Parks, A. L. & Muskavitch, M. A. Delta function is required for bristle organ determination and morphogenesis in *Drosophila*. *Dev. Biol.* **157**, 484–496 (1993).
- Jennings, B., Preiss, A., Delidakis, C. & Bray, S. The Notch signalling pathway is required for Enhancer of split bHLH protein expression during neurogenesis in the *Drosophila* embryo. *Development* **120**, 3537–3548 (1994).
- Lecourtis, M. & Schweisguth, F. The neurogenic suppressor of hairless DNA-binding protein mediates the transcriptional activation of the enhancer of split complex genes triggered by Notch signaling. *Genes Dev.* **9**, 2598–2608 (1995).
- Zeng, C., Younger-Shepherd, S., Jan, L. Y. & Jan, Y. N. Delta and Serrate are redundant Notch ligands required for asymmetric cell divisions within the *Drosophila* sensory organ lineage. *Genes Dev.* **12**, 1086–1091 (1998).
- Qi, H. *et al.* Processing of the notch ligand delta by the metalloprotease Kuzbanian. *Science* **283**, 91–94 (1999).
- Mishra-Gorur, K., Rand, M. D., Perez-Villamil, B. & Artavanis-Tsakonas, S. Down-regulation of Delta by proteolytic processing. *J. Cell Biol.* **159**, 313–324 (2002).
- Kooh, P. J., Fehon, R. G. & Muskavitch, M. A. Implications of dynamic patterns of Delta and Notch expression for cellular interactions during *Drosophila* development. *Development* **117**, 493–507 (1993).
- Sanson, B., Alexandre, C., Fascetti, N. & Vincent, J. P. Engrailed and hedgehog make the range of Wingless asymmetric in *Drosophila* embryos. *Cell* **98**, 207–216 (1999).
- Milan, M., Weihe, U., Perez, L. & Cohen, S. M. The LRR proteins capricious and Tartan mediate cell interactions during DV boundary formation in the *Drosophila* wing. *Cell* **106**, 785–794 (2001).
- Locke, M. & Huie, P. Epidermal feet in insect morphogenesis. *Nature* **293**, 733–735 (1981).
- Fehon, R. G. *et al.* Molecular interactions between the protein products of the neurogenic loci Notch and Delta, two EGF-homologous genes in *Drosophila*. *Cell* **61**, 523–534 (1990).
- Bretscher, A., Edwards, K. & Fehon, R. G. ERM proteins and merlin: integrators at the cell cortex. *Nature Rev. Mol. Cell Biol.* **3**, 586–599 (2002).
- Martin, M. *et al.* Ezrin NH₂-terminal domain inhibits the cell extension activity of the COOH-terminal domain. *J. Cell Biol.* **128**, 1081–1093 (1995).
- Martin, M., Roy, C., Montcourier, P., Sahuquet, A. & Mangeat, P. Three determinants in ezrin are responsible for cell extension activity. *Mol. Biol. Cell* **8**, 1543–1557 (1997).
- Meir, E., von Dassow, G., Munro, E. & Odell, G. M. Robustness, flexibility, and the role of lateral inhibition in the neurogenic network. *Curr. Biol.* **12**, 778–786 (2002).

- Turing, A. M. The chemical basis for morphogenesis. *Phil. Trans. R. Soc. Lond. B* **237**, 37–72 (1952).
- Bellaiche, Y., Gho, M., Kaltschmidt, J. A., Brand, A. H. & Schweisguth, F. Frizzled regulates localization of cell-fate determinants and mitotic spindle rotation during asymmetric cell division. *Nature Cell Biol.* **3**, 50–57 (2001).
- Lee, T. & Luo, L. Mosaic analysis with a repressible cell marker for studies of gene function in neuronal morphogenesis. *Neuron* **22**, 451–461 (1999).
- Hinz, U., Giebel, B. & Campos-Ortega, J. A. The basic-helix-loop-helix domain of *Drosophila* lethal of scute protein is sufficient for proneural function and activates neurogenic genes. *Cell* **76**, 77–87 (1994).
- Seugnet, L., Simpson, P. & Haenlin, M. Transcriptional regulation of Notch and Delta: requirement for neuroblast segregation in *Drosophila*. *Development* **124**, 2015–2025 (1997).
- Haenlin, M., Kunisch, M., Kramatschek, B. & Campos-Ortega, J. A. Genomic regions regulating early embryonic expression of the *Drosophila* neurogenic gene Delta. *Mech. Dev.* **47**, 99–110 (1994).
- Bender, L. B., Kooh, P. J. & Muskavitch, M. A. Complex function and expression of Delta during *Drosophila* oogenesis. *Genetics* **133**, 967–978 (1993).
- Lamb, R. F. *et al.* Essential functions of ezrin in maintenance of cell shape and lamellipodial extension in normal and transformed fibroblasts. *Curr. Biol.* **7**, 682–688 (1997).
- Andreoli, C., Martin, M., Le Borgne, R., Reggio, H. & Mangeat, P. Ezrin has properties to self-associate at the plasma membrane. *J. Cell Sci.* **107**, 2509–2521 (1994).
- Ramirez-Weber, F. A. & Kornberg, T. B. Cytonemes: cellular processes that project to the principal signaling center in *Drosophila* imaginal discs. *Cell* **97**, 599–607 (1999).

Acknowledgements This work was initiated in the unité INSERM 432, Université Montpellier II. We thank S. Baghdiguian, C. Dambly-Chaudière, A. Ghysen, S. Layalle, P.-H. Mangeat and A.-M. Martinez for discussions; C. Dambly-Chaudière, M. Haenlin, F. Schweisguth, J.-P. Vincent, the Bloomington Stock Center and the Developmental Studies Hybridoma Bank for fly stocks and reagents; N. Lautrédu-Audouy for help with confocal microscopy; E. Mérezègue for help with electron microscopy; E. Gazave and A. Sahuquet for help with morphometry and statistics; and C. Roy for critically reading the manuscript. C.d.J. and D.A. thank the researchers into *Drosophila* at the Institut de Génétique Humaine (Montpellier) for hospitality. This work was supported by grants to D.A. from the Association pour la Recherche contre le Cancer and the Université Montpellier II. C.d.J. was supported by the Fondation pour la Recherche Médicale.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to D.A. (daniel@univ-montp2.fr).

Synaptotagmin I is necessary for compensatory synaptic vesicle endocytosis *in vivo*

Kira E. Poskanzer, Kurt W. Marek, Sean T. Sweeney & Graeme W. Davis

Department of Biochemistry and Biophysics, Program in Neuroscience, University of California, San Francisco, 513 Parnassus Avenue, San Francisco, California 94143, USA

Neurotransmission requires a balance of synaptic vesicle exocytosis and endocytosis¹. Synaptotagmin I (Syt I) is widely regarded as the primary calcium sensor for synaptic vesicle exocytosis^{2–6}. Previous biochemical data suggest that Syt I may also function during synaptic vesicle endocytosis^{7–16}; however, ultrastructural analyses at synapses with impaired Syt I function have provided an indirect and conflicting view of the role of Syt I during synaptic vesicle endocytosis^{3,8–10,14}. Until now it has not been possible experimentally to separate the exocytic and endocytic functions of Syt I *in vivo*. Here, we test directly the role of Syt I during endocytosis *in vivo*. We use quantitative live imaging of a pH-sensitive green fluorescent protein fused to a synaptic vesicle protein (synapto-pHluorin) to measure the kinetics of endocytosis in *sytl*-null *Drosophila*. We then combine live imaging of the synapto-pHluorins with photoinactivation of Syt I, through fluorescein-assisted light inactivation, after normal Syt I-mediated vesicle exocytosis. By inactivating Syt I only during endocytosis, we demonstrate that Syt I is necessary for the endocytosis of synaptic vesicles that have undergone exocytosis using a functional Syt I protein.

We tagged the synaptic vesicle protein n-Synaptobrevin (n-Syb) with supercliptic, pH-sensitive green fluorescent protein (GFP)^{17–20}. Expression of n-Syb-synapto-pHluorin (n-Syb-pH) in *Drosophila* neurons does not alter the functional or morphological development of the neuromuscular junction (NMJ). We show that spontaneous and evoked neurotransmitter release, the electrophysiological response to a stimulus train, FM4-64 recycling, and synaptic bouton number are normal in these animals (Supplementary Fig. 1). Quantitative imaging of n-Syb-pH demonstrates a stimulus-locked, reliable increase in fluorescence intensity that decays to baseline with an average time constant ($\tau = 12.6 \pm 1.23$ s) similar to that observed at vertebrate central nervous system synapses (Fig. 1a, b; see also Supplementary Movie 1)^{17,19}.

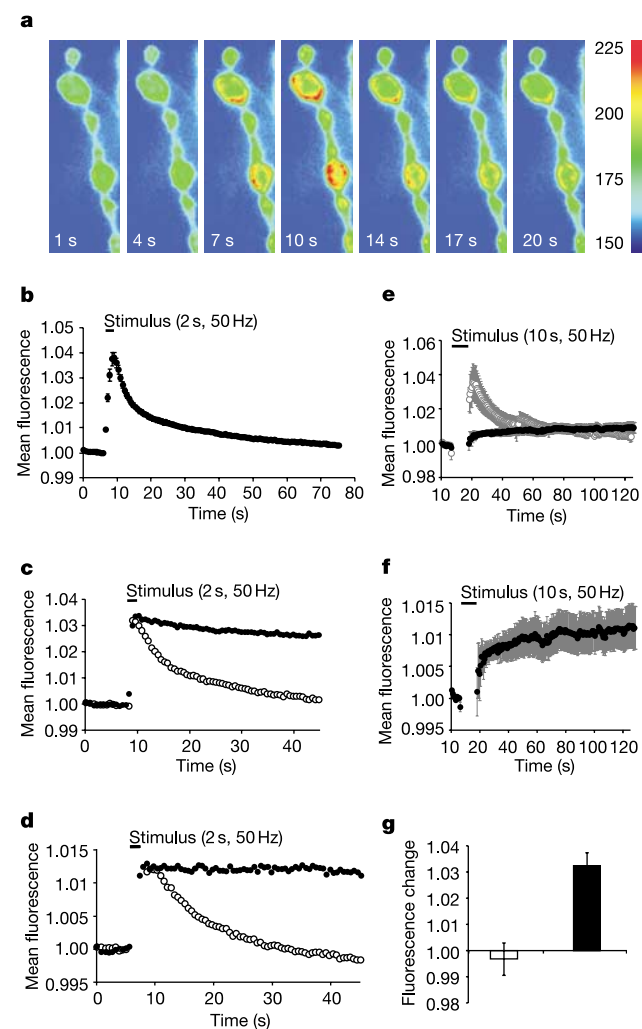


Figure 1 Endocytosis is impaired at *sytI*-null synapses. **a**, n-Syb-pH fluorescence intensity changes at the NMJ before (1 s, 4 s), during (7 s) and after (10–20 s) nerve stimulation. Scale represents arbitrary fluorescence units throughout all panels. **b**, Average n-Syb-pH fluorescence values for the synapse in **a** ($n = 7$ trials). **c**, Representative traces of n-Syb-pH intensity before (open circles) and after (filled circles) addition of bafilomycin (5 μ M, 5 min incubation). **d**, n-Syb-pH intensity in a *shibire*^{ts} mutant at permissive (open circles) and restrictive (filled circles) temperatures. **e**, n-Syb-pH intensity at wild-type (open circles; $n = 4$) and *sytI*-null (filled circles; $n = 7$) synapses. **f**, n-Syb-pH intensity data at *sytI*-null synapses replotted from **e**. **g**, n-Syb-pH fluorescence change in wild-type (open bar; $n = 6$) and *sytI*-null (filled bar; $n = 5$) synapses after a 250-s, 1-Hz stimulus. Fluorescence change is calculated as the difference between pre- and post-stimulus intensities, normalized to pre-stimulus values. No images were taken during the stimulus to minimize photobleaching. *sytI*-null synapses show a significant increase in n-Syb-pH intensity ($P < 0.003$).

Additional experiments demonstrate that n-Syb-pH functions as a reliable pH-sensitive monitor of synaptic vesicle recycling at the *Drosophila* NMJ. First, n-Syb-pH shows robust changes in fluorescence intensity in response to experimental manipulation of pH (Supplementary Fig. 2a, b). Second, bafilomycin, which blocks synaptic vesicle re-acidification, inhibits the fluorescence decay after nerve stimulation, consistent with a failure to re-acidify endocytosed vesicles (Fig. 1c)¹⁸. Third, we generated transgenic flies harbouring n-Syb-pH and a temperature sensitive mutation in dynamin (*shibire*^{ts}). At the permissive temperature, n-Syb-pH fluorescence exhibits a wild-type, stimulus-dependent increase and decay. However, at the restrictive temperature the stimulus-dependent increase in n-Syb-pH fluorescence does not decay, consistent with the complete block of synaptic vesicle endocytosis that has been documented in *shibire*^{ts} (Fig. 1d)²¹. This endocytic block is reversible after shifting to the permissive temperature (data not shown).

To test the function of Syt I during endocytosis, we first used n-Syb-pH to measure the kinetics of endocytosis at the NMJ of *sytI*-null flies. In these experiments the stimulus-dependent (50 Hz for 10 s) increase in n-Syb-pH fluorescence is significantly less in the *sytI*-null synapses compared with controls, consistent with previous electrophysiological data (Fig. 1e, f)^{2,3,22,23}. Furthermore, n-Syb-pH fluorescence after stimulation at the *sytI*-null NMJ does not decay, suggesting a severe impairment of vesicle endocytosis (Fig. 1e, f). If endocytosis is impaired in the *sytI*-null flies then we also expect persistent, low-frequency stimulation to cause a gradual accumulation of n-Syb-pH at the synaptic plasma membrane. Therefore, we stimulated wild-type and *sytI*-null synapses (1 Hz for 250 s) and imaged n-Syb-pH fluorescence before and after the stimulus train. We observed a statistically significant increase in n-Syb-pH fluorescence only at the *sytI*-null synapses, indicating that n-Syb-pH has not been efficiently recycled in the *sytI*-null flies (Fig. 1g). Although these data are consistent with a role for Syt I during vesicle endocytosis, stimulus-evoked exocytosis is severely impaired in the *sytI*-null animals, and we do not know the nature of the Syt I-independent vesicular release that remains.

To test the function of Syt I during endocytosis without affecting Syt I-mediated exocytosis, we pursued experiments to photoinactivate Syt I rapidly after normal stimulus-dependent exocytosis. We have shown previously that fluorescein-assisted light inactivation (FLAsH-FALI) is an efficient and specific method for the photoinactivation of Syt I²⁴. Briefly, Syt I is tagged with a 17-amino-acid tetracycline motif (Syt I4C) that covalently binds the membrane-permeable fluorescein derivative 4',5'-bis(1,3,2-dithioarsolan-2-yl)fluorescein (FLAsH)^{25,26}. Neuronal expression of Syt I4C rescues the *sytI*-null mutation, even when bound to the FLAsH ligand²⁴. Epifluorescent illumination of FLAsH-bound Syt I4C decreases the evoked release in seconds, demonstrating efficient fluorophore-assisted light inactivation (FALI). FLAsH-FALI of Syt I4C does not alter sucrose-evoked release, demonstrating the specificity of Syt I4C photoinactivation²⁴. We have now combined FLAsH-FALI of Syt I4C with the n-Syb-pH technique. To ensure that endogenous Syt I does not affect our results, Syt I4C is always expressed in a *sytI*-null mutant background. As both FLAsH and n-Syb-pH fluoresce in the green spectrum, we took advantage of the fact that epitope-bound FLAsH bleaches three orders of magnitude more rapidly than n-Syb-pH (Supplementary Fig. 2c). Epifluorescent illumination of FLAsH for 2 min decreases the FLAsH signal to near extinction, allowing us to quantitatively image n-Syb-pH.

In the first FLAsH-FALI experiment we used *shibire*^{ts} to block synaptic vesicle endocytosis conditionally and trap vesicles at the plasma membrane that have fused in the presence of functional Syt I4C during stimulation at the restrictive temperature. We then photoinactivated Syt I4C for 2 min, before shifting the preparation to the permissive temperature. If Syt I is necessary for endocytosis, we expect that n-Syb-pH will be retained at the plasma membrane

after the shift to the permissive temperature and the release of the *shibire^{ts}* endocytic block (Fig. 2a). In these experiments n-Syb-pH fluorescence is quantified at the beginning of the shift to the permissive temperature ($t = 0$). A previous ultrastructural study monitored endocytosis at multiple time points in *shibire^{ts}* flies after a shift to the permissive temperature²¹. These previous data showed that endocytosis, after release from a dynamin block, occurs at the order of several minutes, which is slower than that observed after a short stimulus in wild-type flies (see Fig. 1). The slow rate of endocytosis after release from a dynamin block may be caused by the abundance of membrane and vesicle proteins that are trapped at the plasma membrane during the endocytic block.

In this experiment we used transgenic *Drosophila* flies that harbour the *shibire^{ts}* mutation as well as neuronally driven Syt I4C and neuronally driven n-Syb-pH in a *sytI*-null background. These animals were raised at the permissive temperature, and show normal vesicle recycling at this temperature (Supplementary Fig. 2d). In our control condition stimulation at the restrictive temperature in the absence of FAsH induces a stable increase in n-Syb-pH fluorescence that decays upon shift to the permissive temperature, reflecting normal endocytosis (Fig. 2b, c; see also Supplementary Fig. 3, control). This control experiment includes a 2-min illumination before the shift to the permissive temperature to control for the effects of illumination. Our experimental condition includes the presence of FAsH ligand and a 2-min illumination before the shift to the permissive temperature to induce photoinactivation of Syt I4C. In this experiment the n-Syb-pH signal does not decay on shifting to the permissive temperature, indicating that Syt I4C photoinactivation significantly impairs endocytosis (Fig. 2b, c; see also Supplementary Fig. 3, FAsH). To determine whether FAsH fluorescence confounds our assay of endocytosis, we quantified the total synaptic n-Syb-pH fluorescence by the addition of membrane-permeable, high pH NH_4Cl saline (pH 7.4) at the end of each experiment (see Supplementary Fig. 2b). The total n-Syb-pH fluorescence is identical in control and experimental conditions, indicating that FAsH bleaches to near extinction in

these experiments (Fig. 2c). Thus, the lack of n-Syb-pH fluorescence change after Syt I4C photoinactivation cannot be attributed to additional fluorescence contributed by the presence of the FAsH ligand. These results demonstrate that photoinactivation of Syt I4C blocks the endocytosis of synaptic vesicles that have fused with the plasma membrane using a previously functional Syt I4C molecule.

In a second FAsH-FALI experiment we tested whether Syt I4C photoinactivation blocks endocytosis without previously pausing the vesicle cycle using *shibire^{ts}*. Here, we assayed the retention of n-Syb-pH at the synaptic plasma membrane after Syt I4C photoinactivation. FAsH-FALI of Syt I4C was initiated at the time of stimulation (50 Hz for 5 s) and lasted for 3 min (Fig. 3a). If Syt I4C photoinactivation blocks endocytosis that normally occurs during and after the stimulus, then we expect to observe increased fluorescence at the end of the 3-min period due to unrecycled n-Syb-pH retained at the synaptic plasma membrane. We also expect to see a greater decrease in n-Syb-pH fluorescence after an acid wash, which should quench the surface n-Syb-pH fluorescence (Supplementary Fig. 2a)¹⁸. In this protocol, Syt I4C-dependent exocytosis will decline during the stimulation owing to photoinactivation of Syt I4C, so that the total amount of exocytosis will be less in the experimental flies compared with controls²⁴. However, this will cause us to underestimate any potential difference in the accumulation of surface n-Syb-pH caused by Syt I4C photoinactivation

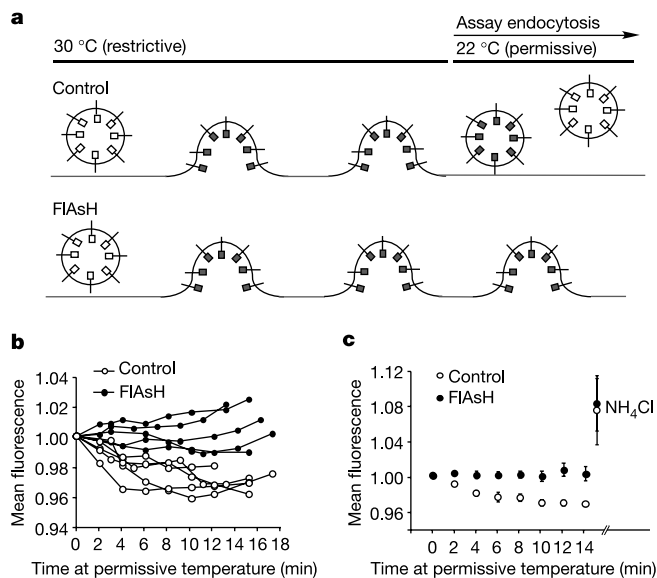


Figure 2 Photoinactivation of Syt I4C impairs endocytosis after normal vesicle fusion. **a**, Schematic of the experimental protocol and results. **b**, Quantification of the n-Syb-pH intensity changes after shift from restrictive to permissive temperatures. Each line represents a single trial from a different animal. Values are normalized to the first time point after shift to the permissive temperature. **c**, Mean values for data in **b**. Data are binned in 2-min intervals. The last point (NH_4Cl) is the intensity after addition of NH_4Cl saline at the conclusion of each trial.

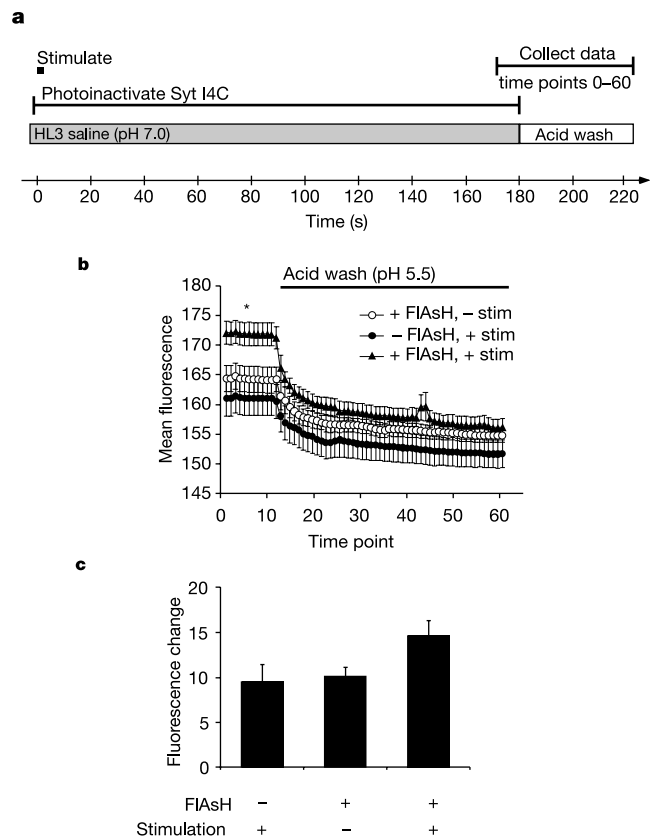


Figure 3 n-Syb-pH remains on the synaptic plasma membrane after FAsH-FALI of Syt I4C. **a**, Protocol for acid wash experiments. **b**, Data shown include the ten time points before addition of the acid wash and the subsequent 50 time points. **c**, Change in intensity from the mean of the first ten time points to the mean of the last ten in **b**. Conditions include the presence (+) or absence (–) of FAsH ligand and presence or absence of stimulation (5 s, 50 Hz). Synapses with Syt I4C photoinactivation and stimulation ($n = 11$) show a significantly greater change in fluorescence intensity than FAsH alone ($n = 9$, $P < 0.01$) or stimulation alone ($n = 12$, $P < 0.01$, t -test). Controls are not significantly different from each other. The genotype used is *syt^{AD4}*, *UAS-syt4C/syt^{D27}*, *UAS-n-Syb-pH/elav^{3E1}*.

compared with controls. As a control, we demonstrated that n-Syb-pH recycles normally in this genetic background (Supplementary Fig. 2d).

In these experiments Syt I4C photoinactivation results in a statistically significant increase in surface n-Syb-pH fluorescence, consistent with impaired endocytosis. We quantified n-Syb-pH fluorescence 10 s before application of the acid wash (Fig. 3a). At this time n-Syb-pH intensity is significantly greater at synapses that have received both stimulation and Syt I4C photoinactivation compared with controls (Fig. 3b, asterisk). Subsequent acid wash reduced n-Syb-pH fluorescence to statistically similar levels in all conditions (Fig. 3b). Therefore, the decrease in surface fluorescence after acid wash is significantly greater at experimental synapses compared with controls (Fig. 3c). These data are consistent with a requirement of Syt I during compensatory endocytosis. Furthermore, these data argue against the possibility that Syt I functions exclusively during vesicle re-acidification, as we have assayed the plasma membrane retention of n-Syb-pH.

Finally, we measured FM4-64 uptake as an alternative assay of vesicle endocytosis. We first demonstrated normal FM4-64 loading and unloading at the *Drosophila* NMJ (Fig. 4b)²⁷. We then took advantage of the fact that endocytosis is slow compared with exocytosis, allowing temporal separation of these processes. First, we stimulated (30 Hz for 1 min) synapses in the presence of FM4-64, causing similar exo- and endocytosis in all conditions. After the stimulus, vesicles will cease exocytosis, but will continue to endocytose, compensating for the large amount of release that occurred during the stimulation. We continued to bathe the synapses in FM4-64 while illuminating for 2 min. This allowed the nerve terminal to endocytose the dye during Syt I4C photoinactivation. After washing the FM4-64 away, we measured the amount of FM4-64 taken up during this endocytic phase of the experiment (Fig. 4a). If Syt I is necessary for compensatory endocytosis then we expect to observe less loading of the nerve terminal in Syt I4C-photoinactivated synapses. FIAH-FALI of Syt I4C results in statistically lower mean fluorescence values of FM4-64 compared with controls, indicating

that Syt I4C photoinactivation impairs endocytosis (Fig. 4c). We observed control levels of FM4-64 loading in the presence of FIAH-bound Syt I4C without any illumination, demonstrating that the ligand-bound Syt I4C itself does not impair endocytosis.

Thus, by temporally uncoupling exo- and endocytosis, we have found an essential role for a known exocytic protein (Syt I) during endocytosis. Although different modes of endocytosis with different time courses may exist^{19,28,29}, a specialized form of clathrin-mediated endocytosis, termed compensatory endocytosis, participates in the maintenance of the synaptic vesicle pool^{1,19}. Compensatory endocytosis involves the formation of synaptic vesicles that are surrounded by clathrin and the AP-2 adaptor complex, as well as associated adaptor proteins¹. Our data are consistent with a model based on previous biochemical data suggesting that Syt I nucleates the formation of clathrin-coated vesicles by serving as a synaptic plasma membrane docking site for the AP-2 complex^{7,11–13,15,16}. This model provides a means to link the extent of synaptic vesicle exocytosis and endocytosis. Finally, it has been demonstrated that multiple Syt proteins can bind AP-2 (ref. 30). However, as we have disrupted acutely Syt I, there should not be an opportunity for other Syt proteins to bind AP-2 and nucleate endocytosis in our experiments. □

Methods

Construction of UAS-n-Syb-pH

Superecliptic pHluorin GFP was obtained from D. DeAngelis and J. Rothman. To fuse the pHluorin GFP to the carboxy terminus of n-Syb, a BsrB1 site was added to the 3' end of an n-Syb complementary DNA (from T. Schwarz) using the oligonucleotides 5' n-Syb (ACAGCCGAATTCGCTGAGGC) and 3' n-Syb-BsrB1 (CCAATTAGATCTCCGCTCA CGCCGCGGTGATCGCC). An *Xba*I site was added to the 3' end and a BsrB1 site was added to the 5' end of the pHluorin GFP using the oligonucleotides 5' GFP-BsrB1 (CACGGCGCGTGGAGCGCGGAGCGCGGG) and 3' GFP-Xba (TAGCTTCTAG ATTAACCGGTTTGTATAG). Polymerase chain reaction fragments were then digested with BsrB1 and ligated together. The appropriate fragment was purified and used as a template for a second round of polymerase chain reaction using the oligonucleotides 5' n-Syb and 3' GFP-Xba. The resulting product was then digested with *Eco*R1 and *Xba*I and ligated into pUAST. The resulting clone was sequenced before injecting into *yw* embryos.

Flies

The UAS-n-Syb-pH transgene was recombined with *elav*^{V3E1}-*GAL4*, and balanced with a TM6b chromosome. For the *shibire*^{ts} experiments, *w*⁻, *shi*^{ts2} was crossed to UAS-n-Syb-pH, *elav*^{V3E1}/TM6b and male larvae were selected against TM6b. The *sytl*-null flies were made by crossing *sytl*^{AD4}/CyO-GFP; UAS-n-Syb-pH/TM6b to *sytl*^{D27}/CyO-GFP; *elav*^{V3E1}/TM6b and selecting against TM6b and CyO-GFP expression. For the FIAH-FALI experiments, *sytl*^{AD4}, UAS-syt4C/CyO-GFP; UAS-n-Syb-pH/TM6b was crossed to *sytl*^{D27}/CyO-GFP; *elav*^{V3E1}/TM6b and larvae were selected against CyO-GFP expression. The FIAH-FALI *shibire*^{ts} experiments were performed by crossing *w*⁻, *shi*^{ts2}; *sytl*^{D27}/CyO-GFP; *elav*^{V3E1}/TM6b to *sytl*^{AD4}, UAS-syt4C/CyO-GFP; UAS-n-Syb-pH/TM6b and selecting males against CyO-GFP and TM6b expression.

Solutions

Third instar *Drosophila* larvae were dissected in Ca^{2+} -free HL3 saline, supplemented with calcium and adjusted to pH 7.0 (ref. 24). All experiments were performed in 0.5 mM Ca^{2+} with the exception of experiments in Fig. 1e–g, which were performed in 2.0 mM Ca^{2+} . For FIAH experiments, dissected animals were either incubated in FIAH solution (Invitrogen) or in a mock-FIAH solution, then rinsed to remove unbound FIAH²⁴. Larvae were transferred back to HL3 saline before imaging. HL3 saline was adjusted to pH 5.5 in the acid wash experiments. Ammonium chloride (NH_4Cl) HL3 saline was made by replacing 50 mM NaCl with NH_4Cl and adjusting to pH 7.4.

Imaging and analysis

Images were captured on a Zeiss Axioskop 2 microscope with a water-immersion objective ($\times 100$, 0.97 N.A.) using SlideBook software (Intelligent Imaging Innovations). Images were acquired with a Quantix cooled CCD camera including a back-thinned EEV57 chip (Roper Scientific). Illumination was provided by a 175 W xenon arc lamp and GFP or CY3 filter sets. Time-lapse imaging (500 ms exposures) was used to measure the fast time course of fluorescence changes (Figs 1, 3). Data were quantified by defining a region of interest including multiple synaptic boutons (each bouton contains multiple active zones) using the threshold function of SlideBook. The mean fluorescence intensity of this region was calculated for each time point. FIAH-FALI was performed by illuminating the preparations as for fluorescence imaging. The NMJ at muscles 6 and 7 in segments A2–A4 were selected in all experiments. In all experiments scale represents arbitrary fluorescence units. Data are expressed as average $\pm$ standard error of the mean.

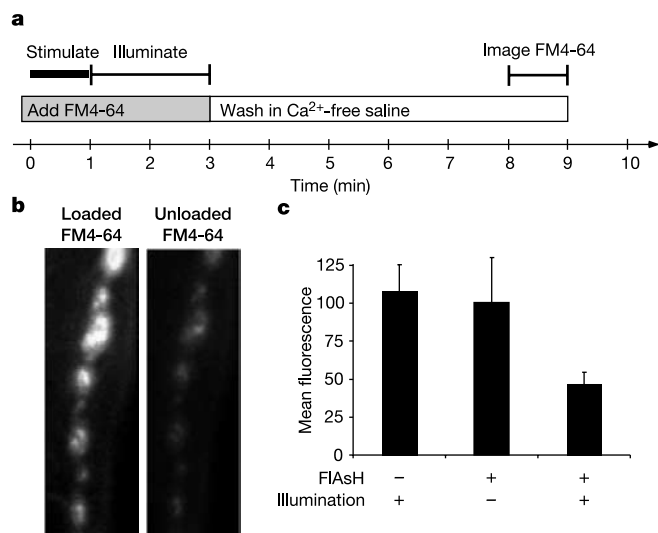


Figure 4 Photoinactivation of Syt I4C impairs stimulus-dependent FM4-64 loading at the NMJ. **a**, Protocol for FM4-64 experiments. **b**, Images demonstrating efficient loading (left) and unloading (right) of FM4-64. **c**, Quantification of FM4-64 loading at synapses in the presence (+) or absence (-) of FIAH ligand and in the presence or absence of illumination. Loading of FM4-64 is significantly less ($P < 0.02$) at synapses that have undergone photoinactivation of Syt I4C (plus FIAH, plus illumination, $n = 6$) compared with controls (minus FIAH, plus illumination, $n = 7$; plus FIAH, minus illumination, $n = 8$).

Acid wash experiments

Synapses were stimulated at the segmental nerve for 5 s at 50 Hz in 2.0 mM Ca^{2+} and imaged for 3 min before the HL3 saline (pH 7.0) in the bath was replaced with low pH acid wash (pH 5.5).

FM4-64 experiments

A total of 10 μM FM4-64 (Molecular Probes) was made in HL3 plus 0.5 mM Ca^{2+} saline and added to the bath. The nerve was stimulated for 1 min at 30 Hz. The synapse was then illuminated for 2 min. The preparation was washed repeatedly in Ca^{2+} -free HL3 for 5 min, and an image of the synapse was captured. To analyse FM4-64 uptake the n-Syb-pH fluorescence in the green channel was used to define the region of interest, and the mean intensity of the fluorescence in the red channel defined by this region was calculated. We selected a small region of interest over the muscle background and the mean intensity of this region was subtracted from the synapse intensity to correct for variability in background FM4-64 levels.

shibire^{ts} experiments

Temperature was controlled by a heating/cooling stage with a Peltier module (Brook Industries) and monitored with a temperature probe in the imaging chamber. In Fig. 1d the nerve was stimulated identically (2 s, 50 Hz) at the permissive (22 °C) and restrictive temperatures (30 °C). In the FIAsh experiments (Fig. 2) we stimulated at the restrictive temperature (2 s, 50 Hz, in HL3) every 10 s for 200 s to deplete synaptic vesicles. After stimulation, the synapse was illuminated for 2 min. The chamber was then brought down to the permissive temperature. Images of the synapse were taken at about 2 min intervals after return to the permissive temperature, and the mean intensity of the synapse was measured at each time point.

Received 25 July; accepted 6 November 2003; doi:10.1038/nature02184.

Published online 23 November 2003.

- Jarousse, N. & Kelly, R. B. Endocytotic mechanisms in synapses. *Curr. Opin. Cell Biol.* **13**, 461–469 (2001).
- Fernandez-Chacon, R. *et al.* Synaptotagmin I functions as a calcium regulator of release probability. *Nature* **410**, 41–49 (2001).
- Geppert, M. *et al.* Synaptotagmin I: a major Ca^{2+} sensor for transmitter release at a central synapse. *Cell* **79**, 717–727 (1994).
- Littleton, J. T., Stern, M., Schulze, K., Perin, M. & Bellen, H. J. Mutational analysis of *Drosophila* synaptotagmin demonstrates its essential role in Ca^{2+} -activated neurotransmitter release. *Cell* **74**, 1125–1134 (1993).
- Perin, M. S., Fried, V. A., Mignery, G. A., Jahn, R. & Sudhof, T. C. Phospholipid binding by a synaptic vesicle protein homologous to the regulatory region of protein kinase C. *Nature* **345**, 260–263 (1990).
- Brose, N., Petrenko, A. G., Sudhof, T. C. & Jahn, R. Synaptotagmin: a calcium sensor on the synaptic vesicle surface. *Science* **256**, 1021–1025 (1992).
- Fergestad, T. & Brodie, K. Interaction of stoned and synaptotagmin in synaptic vesicle endocytosis. *J. Neurosci.* **21**, 1218–1227 (2001).
- Jorgensen, E. M. *et al.* Defective recycling of synaptic vesicles in synaptotagmin mutants of *Caenorhabditis elegans*. *Nature* **378**, 196–199 (1995).
- Littleton, J. T. *et al.* Synaptotagmin mutants reveal essential functions for the C2B domain in Ca^{2+} -triggered fusion and recycling of synaptic vesicles *in vivo*. *J. Neurosci.* **21**, 1421–1433 (2001).
- Reist, N. E. *et al.* Morphologically docked synaptic vesicles are reduced in synaptotagmin mutants of *Drosophila*. *J. Neurosci.* **18**, 7662–7673 (1998).
- Jarousse, N. & Kelly, R. B. The AP2 binding site of synaptotagmin 1 is not an internalization signal but a regulator of endocytosis. *J. Cell Biol.* **154**, 857–866 (2001).
- Zhang, J. Z., Davletov, B. A., Sudhof, T. C. & Anderson, R. G. Synaptotagmin I is a high affinity receptor for clathrin AP-2: implications for membrane recycling. *Cell* **78**, 751–760 (1994).
- Hauke, V. & De Camilli, P. AP-2 recruitment to synaptotagmin stimulated by tyrosine-based endocytic motifs. *Science* **285**, 1268–1271 (1999).
- Fukuda, M. *et al.* Role of the C2B domain of synaptotagmin in vesicular release and recycling as determined by specific antibody injection into the squid giant synapse preterminal. *Proc. Natl Acad. Sci. USA* **92**, 10708–10712 (1995).
- Hauke, V., Wenk, M. R., Chapman, E. R., Farsad, K. & De Camilli, P. Dual interaction of synaptotagmin with mu2- and alpha-adaptin facilitates clathrin-coated pit nucleation. *EMBO J.* **19**, 6011–6019 (2000).
- von Poser, C. *et al.* Synaptotagmin regulation of coated pit assembly. *J. Biol. Chem.* **275**, 30916–30924 (2000).
- Sankaranarayanan, S. & Ryan, T. A. Real-time measurements of vesicle-SNARE recycling in synapses of the central nervous system. *Nature Cell Biol.* **2**, 197–204 (2000).
- Sankaranarayanan, S. & Ryan, T. A. Calcium accelerates endocytosis of vSNAREs at hippocampal synapses. *Nature Neurosci.* **4**, 129–136 (2001).
- Gandhi, S. P. & Stevens, C. F. Three modes of synaptic vesicular recycling revealed by single-vesicle imaging. *Nature* **423**, 607–613 (2003).
- Miesenböck, G., De Angelis, D. A. & Rothman, J. E. Visualizing secretion and synaptic transmission with pH-sensitive green fluorescent proteins. *Nature* **394**, 192–195 (1998).
- Koenig, J. H. & Ikeda, K. Synaptic vesicles have two distinct recycling pathways. *J. Cell Biol.* **135**, 797–808 (1996).
- Loewen, C. A., Mackler, J. M. & Reist, N. E. *Drosophila* synaptotagmin I null mutants survive to early adulthood. *Genesis* **31**, 30–36 (2001).
- Yoshihara, M. & Littleton, J. T. Synaptotagmin I functions as a calcium sensor to synchronize neurotransmitter release. *Neuron* **36**, 897–908 (2002).
- Marek, K. W. & Davis, G. W. Transgenically encoded protein photoinactivation (FIAsh-FALI): acute inactivation of synaptotagmin I. *Neuron* **36**, 805–813 (2002).
- Griffin, B. A., Adams, S. R. & Tsien, R. Y. Specific covalent labeling of recombinant protein molecules inside live cells. *Science* **281**, 269–272 (1998).

- Gaietta, G. *et al.* Multicolor and electron microscopic imaging of connexin trafficking. *Science* **296**, 503–507 (2002).
- Kuromi, H. & Kidokoro, Y. Two distinct pools of synaptic vesicles in single presynaptic boutons in a temperature-sensitive *Drosophila* mutant, shibire. *Neuron* **20**, 917–925 (1998).
- Aravanis, A. M., Pyle, J. L. & Tsien, R. W. Single synaptic vesicles fusing transiently and successively without loss of identity. *Nature* **423**, 643–647 (2003).
- Verstreken, P. *et al.* Endophilin mutations block clathrin-mediated endocytosis but not neurotransmitter release. *Cell* **109**, 101–112 (2002).
- Li, C. *et al.* Ca^{2+} -dependent and -independent activities of neural and non-neural synaptotagmins. *Nature* **375**, 594–599 (1995).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank M. Ramaswami and I. Robinson for *Drosophila* stocks, and D. DeAngelis, J. Rothman and R. Kelly for the superecliptic pHluorin GFP construct. We also thank E. Heckscher for comments on an earlier version of the manuscript.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to G.W.D. (gdavis@biochem.ucsf.edu).

Lipid packing sensed by ArfGAP1 couples COPI coat disassembly to membrane bilayer curvature

Joëlle Bigay¹, Pierre Gounon², Sylviane Robineau¹ & Bruno Antonny¹

¹Institut de Pharmacologie Moléculaire et Cellulaire, CNRS, 660 route des Lucioles, 06560 Valbonne-Sophia-Antipolis, France

²Centre Commun de Microscopie Appliquée, Université de Nice, Parc Valrose, 06103 Nice cedex 2, France

Protein coats deform flat lipid membranes into buds and capture membrane proteins to form transport vesicles^{1–3}. The assembly/disassembly cycle of the COPI coat on Golgi membranes is coupled to the GTP/GDP cycle of the small G protein Arf1. At the heart of this coupling is the specific interaction of membrane-bound Arf1–GTP with coatamer, a complex of seven proteins that forms the building unit of the COPI coat^{4–7}. Although COPI coat disassembly requires the catalysis of GTP hydrolysis in Arf1 by a specific GTPase-activating protein (ArfGAP1)^{8–10}, the precise timing of this reaction during COPI vesicle formation is not known. Using time-resolved assays for COPI dynamics on liposomes of controlled size, we show that the rate of ArfGAP1-catalysed GTP hydrolysis in Arf1 and the rate of COPI disassembly increase over two orders of magnitude as the curvature of the lipid bilayer increases and approaches that of a typical transport vesicle. This leads to a model for COPI dynamics in which GTP hydrolysis in Arf1 is organized temporally and spatially according to the changes in lipid packing induced by the coat.

We reported previously that a fragment of ArfGAP1 (amino acids 1–257), which includes the catalytic zinc-finger domain, and Gcs1p, a yeast ArfGAP1 homologue, bind preferentially to liposomes containing conical lipids versus cylindrical lipids¹¹. The binding increased gradually as the size of lipid polar heads decreased and as the cross-section of the acyl chains increased. Because the marked preference of [1–257]ArfGAP1 for conical lipids correlated with its activity on liposome-bound Arf1–GTP¹¹, we suggested that a loose lipid packing favours the positioning of ArfGAP1 towards membrane-bound Arf1–GTP. Similar effects of the balance between conical and cylindrical lipids were observed

on the binding and the activity of full-length ArfGAP1 (data not shown).

Because the link between COPI vesicle biogenesis and lipid metabolism remains elusive^{12,13}, we reasoned that if lipid packing was at the root cause of the effects observed, changes in GAP activity should be observed by varying the liposome radius while keeping the lipid composition constant. This seemed promising in the context of COPI vesicle formation, because of the membrane distortions imposed by the coat during budding and notably the increase in lipid spacing in the coat-facing leaflet. Liposomes with a composition approaching that of Golgi membranes were prepared by sequential extrusions through 0.4, 0.2, 0.1 and 0.03 μm (pore size) polycarbonate filters¹⁴. We checked, using thin-layer chromatography and a phosphorus assay (Fig. 1a and data not shown), that the extrusion steps did not affect the lipid concentration and the composition of the liposomes. We assessed the size distribution of the liposomes by dynamic light scattering (Fig. 1b). To monitor the activity of full-length ArfGAP1 on liposome-bound Arf1-GTP, we took advantage of the large tryptophan fluorescence change that correlates with the transition of Arf1 between GTP- and GDP-bound forms¹¹. Strikingly, whereas liposome radius had no influence on Arf1 activation, a dramatic effect on the kinetics of Arf1 inactivation was observed (Fig. 1c; see also Supplementary Information). With large liposomes (hydrodynamic radius,

$R_h = 148 \pm 50$ nm), the half-time of Arf1 inactivation initiated by the addition of 25 nM full-length ArfGAP1 was 306 s. As the liposome radius decreased, the half-time decreased to reach a value of 8 s with small liposomes ($R_h = 38 \pm 15$ nm; Fig. 1d). Interestingly, the sensitivity to membrane curvature became very steep at radii approaching that of typical coated vesicles (<50 nm). Because the ratio between ArfGAP1 and Arf in these experiments was 1:20, the rate constant for GTP hydrolysis on small ($R_h = 35\text{--}40$ nm) liposomes was in the subsecond range ($k_{\text{cat}} > 1 \text{ s}^{-1}$).

The function of Arf proteins is not restricted to coated-vesicle formation and it was important to determine whether other Arf-GAP proteins, which are not involved in membrane trafficking, are sensitive to membrane curvature. We used a construct named PZA, derived from ASAP1, a protein involved in actin rearrangements^{15,16}. PZA consists of a PH domain, which binds phosphatidylinositol bisphosphate, a zinc-finger 'GAP' domain homologous to the catalytic domain of ArfGAP1, and ankyrin repeats. Although PZA-catalysed GTP hydrolysis in Arf1 clearly occurred on the liposome surface, as demonstrated by the effect of phosphatidylinositol bisphosphate, the activity of PZA was insensitive to liposome radius (Fig. 1e, f).

Next, we wished to determine whether membrane curvature directly influences the rate of COPI coat disassembly. Previous biochemical and electron microscopy studies have demonstrated that stable COPI-like coated vesicles can form when large liposomes are incubated with Arf1, coatamer and a non-hydrolysable GTP analogue^{17,18}. To monitor the assembly/disassembly cycle of the COPI coat in this minimal system, we used a light-scattering assay developed to study the dynamics of the COPII coat¹⁹. Briefly, the intensity of the light scattered by liposomes increased as the small G protein and the COP complexes are recruited to the lipid surface. Conversely, when the proteins dissociate, the intensity diminishes. In a first set of experiments, we used relatively large liposomes ($R_h = 150$ nm) containing 2 mol% of a lipopeptide that imitates the cytosolic tail of the membrane protein p23 and favours coatamer recruitment¹⁸. In Fig. 2a, the order of additions follows the established order of events in COPI assembly and disassembly. Arf1-GDP was first activated by the addition of GTP and by lowering temporarily the free Mg^{2+} concentration from 1 mM to 1 μM with EDTA. During this stage, a small light-scattering increase was observed, which correlated with the GDP/GTP exchange reaction as monitored by tryptophan fluorescence (data not shown). The subsequent addition of coatamer induced a large and instantaneous light-scattering increase. Last, upon the addition of a catalytic amount of ArfGAP1 the signal returned to the initial level within a few minutes. Inverting the order of protein additions (Fig. 2b) or replacing GTP by GTP γS , a non-hydrolysable analogue (Fig. 2c), demonstrated that these three signals correspond respectively to Arf recruitment, coatamer recruitment and ArfGAP1-induced COPI disassembly.

The rate of ArfGAP1-induced COPI disassembly as monitored by light-scattering (Fig. 2a, continuous trace) was faster than the rate of ArfGAP1-catalysed GTP hydrolysis on isolated Arf1, as monitored by fluorescence (Fig. 2a, dashed trace). Several factors could account for this difference. First, coatamer and p23 proteins could modulate ArfGAP1 activity²⁰. In addition, COPI-induced membrane bending could increase ArfGAP1 activity. We visualized, using electron microscopy, the membrane deformations induced by the COPI coat before the addition of ArfGAP1 (Fig. 2d). In agreement with ref. 18, COPI-coated vesicles clearly bud off from some liposomes. Yet, COPI-coated membranes with a clear budding profile and small (<100 nm) COPI-like coated vesicles constituted no more than 15% of the total surface of COPI-coated membranes. Therefore the rate of COPI coat disassembly measured in Fig. 2a essentially reflects the dynamics of weakly curved coated profiles.

To determine unambiguously the effect of membrane curvature

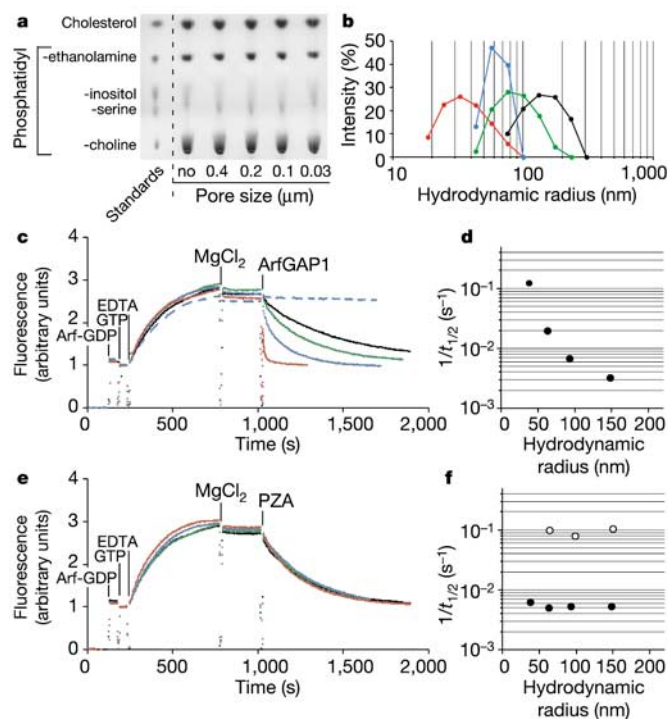


Figure 1 Membrane curvature increases ArfGAP1 activity. **a**, Thin-layer chromatography of Golgi-like liposomes prepared by sequential extrusion through polycarbonate filters of decreasing pore size. **b**, Radius distribution of the extruded liposomes as determined by dynamic light scattering. Pore size in μm : 0.4 (black), 0.2 (green), 0.1 (blue) and 0.03 (red). **c**, Tryptophan fluorescence measurements of the GDP/GTP cycle of Arf1 (0.5 μM) on the liposomes characterized in **a** and **b**. Arf1 was activated by the addition of GTP (100 μM) and by lowering the concentration of Mg^{2+} with EDTA. Inactivation was initiated by the addition of ArfGAP1 (25 nM). Dashed trace: control with GTP γS . **d**, Rate ($1/t_{1/2}$) of ArfGAP1-catalysed GTP hydrolysis as a function of the mean hydrodynamic radius (R_h). **e**, **f**, Control experiments with PZA (25 nM). Time courses and black circles represent standard liposomes; white circles represent standard liposomes supplemented with 2% phosphatidylinositol bisphosphate.

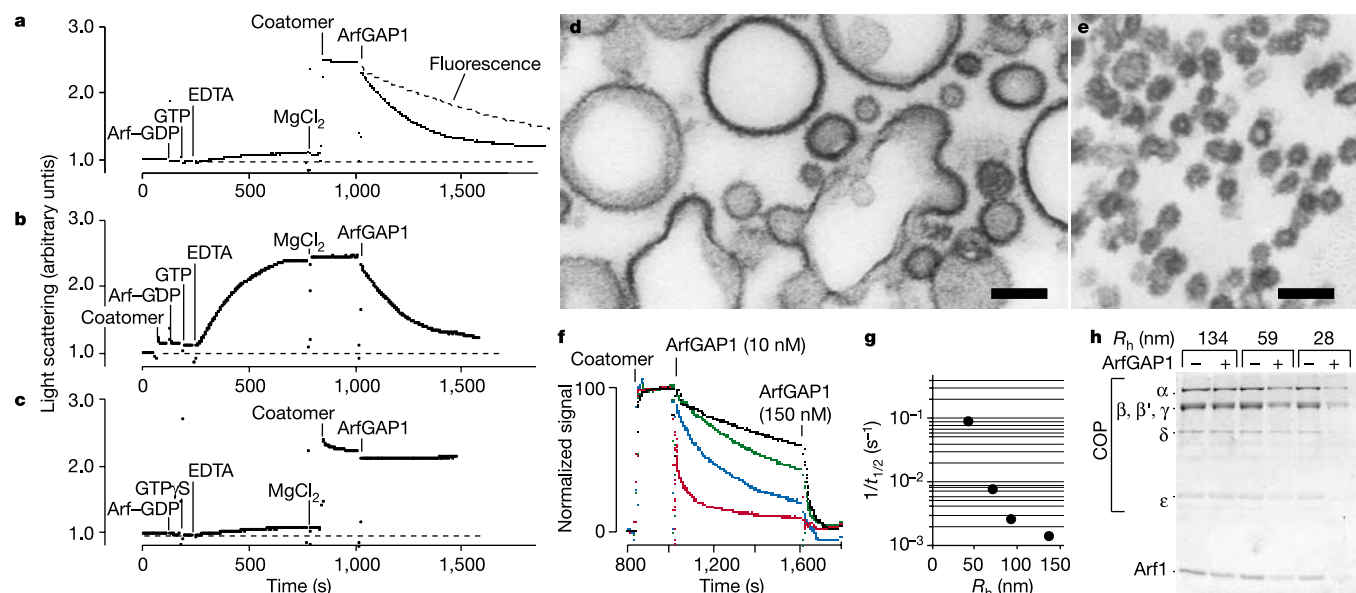


Figure 2 Membrane curvature increases the rate of COPI disassembly. **a–c**, Assembly–disassembly cycle of the COPI coat on large liposomes ($R_h = 150$ nm; p23 lipopeptide: 2%). Arf1–GDP (0.5 μ M), nucleotide (100 μ M), coatmer (0.15 μ M) and ArfGAP1 (25 nM) were added to the liposome suspension as indicated. The scattering of light was measured continuously. The dashed trace reports the time course of Arf1 inactivation as measured by tryptophan fluorescence in the absence of coatmer and lipopeptide.

d, e, Coated profiles observed by electron microscopy on large (**d**, $R_h = 145$ nm) or small (**e**, $R_h = 31$ nm) liposomes after incubation with Arf1, GTP and coatmer before the addition of ArfGAP1. Scale bar, 100 nm. **f**, Same as in **a** with liposomes of decreasing radius (same colour code as in Fig. 1). **g**, Rate ($1/t_{1/2}$) of COPI disassembly as a function of liposome hydrodynamic radius. **h**, Sedimentation analysis of liposome-bound proteins before (–) and 1 min after (+) the addition of 10 nM ArfGAP1.

on COPI dynamics, the light-scattering assay shown in Fig. 2a was repeated with liposomes of decreasing radius. Figure 2e shows an example of small ($R_h = 31$ nm) liposomes covered with the COPI coat. Figure 2f and g shows that the rate of COPI coat disassembly increased over two orders of magnitude as the liposome radius decreased. Thus, whereas the COPI coat associated with large liposomes ($R_h = 140$ nm) was quite resistant to ArfGAP1 ($t_{1/2} = 700$ s), it dissociated from small liposomes ($R_h = 35$ nm) within a few seconds ($t_{1/2} = 11$ s) once a catalytic amount of ArfGAP1 was added. Again, the sensitivity to membrane curvature was very high at low radii (Fig. 2g). These findings were

confirmed by sedimentation experiments in which liposome-bound proteins were analysed before and after the addition of ArfGAP1 (Fig. 2h).

Ionic complexes between fluoride and aluminium (AlF_x) can imitate a γ -phosphate group en route for hydrolysis²¹. For small G proteins, the binding of AlF_x requires the catalytic machinery provided by a GAP protein²². AlF_x promotes the formation of COPI-coated-vesicles, but the molecular basis for this effect is not known^{4,23}. We reasoned that if the effect of AlF_x on COPI relies on the stabilization of an interaction of Arf1–GDP with AlF_x and ArfGAP1, and if this interaction depends on membrane curvature, one may observe an effect of AlF_x on reconstituted COPI reactions performed with small liposomes. Coatmer, ArfGAP1 and Arf1–GDP were added to liposomes of decreasing radius in the presence or in the absence of AlF_x . At the end of the experiments, GTP γ S and

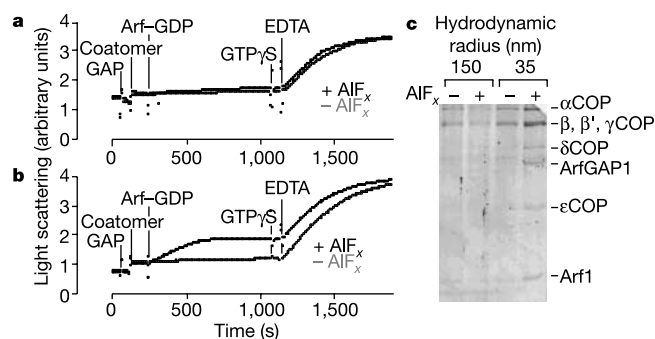


Figure 3 Membrane curvature facilitates the co-assembly of Arf1–GDP with ArfGAP1 and coatmer in the presence of AlF_x . **a**, Light-scattering measurements on large liposomes ($R_h = 150$ nm) with (black trace) or without (grey trace) 8 mM KF and 400 μ M AlCl_3 . ArfGAP1 (0.5 μ M), coatmer (0.3 μ M) and Arf1–GDP (1 μ M) were sequentially added. **b**, Same as in **a** with small liposomes ($R_h = 35$ nm). **c**, Sedimentation analysis. Experiments similar to that shown in **a** and **b** were performed except that no GTP γ S was added. At time = 1,100 s, the samples were centrifuged and the pellets were analysed by SDS–PAGE.

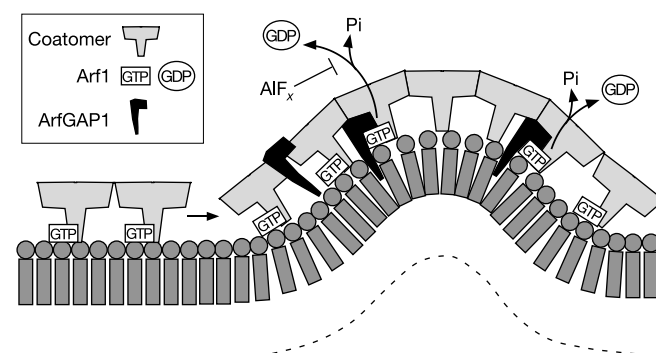


Figure 4 Model for the temporal and spatial distribution of ArfGAP1-catalysed GTP hydrolysis in Arf1 during the formation of a COPI bud. For simplicity, the interaction of the COPI machinery with membrane proteins, which could also influence the rate of GTP hydrolysis, is not shown. Pi, inorganic phosphate.

EDTA were added to provide a reference for the maximal binding signal. On large liposomes supplemented with 2% p23 lipopeptide, the light-scattering recordings with or without Arf_x were almost indistinguishable and no binding signal was observed before the addition of GTP γ S (Fig. 3a, c). In contrast, an Arf_x-dependent signal was detected on small liposomes (Fig. 3b, c). Importantly, whatever the order of protein additions used, the Arf_x-dependent signal was observed after the third protein addition (for example, Arf1-GDP in Fig. 3b), suggesting that Arf_x stabilizes a ternary complex between Arf1-GDP, ArfGAP1 and coatamer provided that the supporting bilayer is sufficiently bent. We note that Arf1 concentrates in buds compared with flat Golgi cisternae in the presence of Arf_x (ref. 4). In conclusion, the effect of Arf_x suggests that the adjustment of the enzymatic machinery that drives GTP hydrolysis in COPI is governed at a distance by the penetration of ArfGAP1 into the coat-facing leaflet as the membrane becomes curved.

The remarkable sensitivity of ArfGAP1 to membrane curvature determines a spatial and temporal programme for GTP hydrolysis in a COPI bud (Fig. 4). As a polymerized COPI coat increases membrane curvature, the number of Arf1-GTP molecules that hold the coat should decrease dramatically. In line with what has been proposed for other coats^{24,25}, this decrease is probably compensated by an increase in the lateral interactions between COP proteins as the coat lattice becomes spherical. Furthermore, because membrane curvature at the periphery of the coated area is negative, a ring of Arf1-GTP molecules should be protected from ArfGAP1 and may keep the coat in a metastable state, ready to disassemble as soon as membrane fission occurs (when membrane curvature becomes entirely positive). Overall, such a mechanism would explain why the release of Arf1 from Golgi membranes is faster than the release of coatamer and that vesicles with an unstable coat and with a low Arf1/coatamer ratio form in the presence of ArfGAP1 and GTP^{26,27}. □

Methods

Proteins

Arf1 was coexpressed in *Escherichia coli* with N-myristoyl-transferase and the myristoylated form was purified as described²⁸. Rabbit-liver coatamer was purified as described²⁹ and further purified by gel-filtration on a Sephacryl S-300 column. Full length ArfGAP1 with a carboxy-terminal hexahistidine tag was purified from Sf9 cells by nickel and mono-Q chromatography.

Liposomes

Liposomes were prepared by extrusion through polycarbonate filters¹⁴. The composition was 50 mol% egg phosphatidylcholine, 19% liver phosphatidylethanolamine, 5% brain phosphatidylserine, 10% liver phosphatidylinositol and 16% cholesterol. When indicated, 2% phosphatidylinositol bisphosphate was added at the expense of phosphatidylinositol. To check that the lipid composition remained constant throughout the sequential extrusion protocol, the liposomes were analysed by thin-layer chromatography using a chloroform/methanol/water/acetic acid (60:50:4:1) solvent system. Lipids were detected by iodine. The p23 lipopeptide was synthesized and purified as described²⁹ and added from a stock solution in DMSO to the extruded liposomes to give a final surface concentration of 2 mol lipopeptide/100 mol exposed lipid.

Dynamic light scattering

The hydrodynamic radius of the extruded liposomes was assessed by dynamic light scattering using a Dynapro MSX instrument at a final lipid concentration of 0.1 mM.

Tryptophan fluorescence and static light-scattering measurements

All experiments were performed at 37 °C. The buffer used was HEPES 50 mM, pH 7.2, potassium acetate 120 mM, MgCl₂ 1 mM, DTT 1 mM. The final lipid concentration was 0.2 mM (fluorescence) or 0.1 mM (light scattering). Tryptophan fluorescence (excitation, 297.5 nm; emission, 340 nm) and light scattering (350 nm) were measured in a small quartz cuvette (volume, 100 μ l) at right angles. The kinetics of ArfGAP1-catalysed GTP hydrolysis on Arf1 and the kinetics of COPI disassembly could not be fitted by a single exponential function (this probably reflects the heterogeneity in size of the liposomes). We determined graphically the apparent half-time ($t_{1/2}$) and used $1/t_{1/2}$ as a rate unit.

Sedimentation

Liposomes and proteins were incubated in small polycarbonate tubes under the same conditions as that used for the light-scattering experiments. The samples were centrifuged

at 50,000g for 15 min at 25 °C and the pellets were analysed by SDS-polyacrylamide gel electrophoresis (PAGE) with Sypro orange staining.

Electron microscopy

Samples from light scattering experiments were collected, fixed with 2% glutaraldehyde, centrifuged at 100,000g for 30 min and the pellet was then processed for electron microscopy as described³⁰.

Received 27 June; accepted 7 October 2003; doi:10.1038/nature02108.

- Rothman, J. E. & Wieland, F. T. Protein sorting by transport vesicles. *Science* **272**, 227–234 (1996).
- Schekman, R. & Orci, L. Coat proteins and vesicle budding. *Science* **271**, 1526–1533 (1996).
- Kirchhausen, T. Three ways to make a vesicle. *Nature Rev. Mol. Cell Biol.* **1**, 187–198 (2000).
- Serafini, T. *et al.* ADP-ribosylation factor is a subunit of the coat of Golgi-derived COP-coated vesicles: a novel role for a GTP-binding protein. *Cell* **67**, 239–253 (1991).
- Donaldson, J. G., Cassel, D., Kahn, R. A. & Klausner, R. D. ADP-ribosylation factor, a small GTP-binding protein, is required for binding of the coatamer protein beta-COP to Golgi membranes. *Proc. Natl Acad. Sci. USA* **89**, 6408–6412 (1992).
- Orci, L., Palmer, D. J., Amherdt, M. & Rothman, J. E. Coated vesicle assembly in the Golgi requires only coatamer and ARF proteins from the cytosol. *Nature* **364**, 732–734 (1993).
- Zhao, L. *et al.* Direct and GTP-dependent interaction of ADP ribosylation factor 1 with coatamer subunit beta. *Proc. Natl Acad. Sci. USA* **94**, 4418–4423 (1997).
- Tanigawa, G. *et al.* Hydrolysis of bound GTP by ARF protein triggers uncoating of Golgi-derived COP-coated vesicles. *J. Cell Biol.* **123**, 1365–1371 (1993).
- Cukierman, E., Huber, I., Rotman, M. & Cassel, D. The ARF1 GTPase-activating protein: zinc finger motif and Golgi complex localization. *Science* **270**, 1999–2002 (1995).
- Reinhard, C., Schweikert, M., Wieland, F. T. & Nickel, W. Functional reconstitution of COPI coat assembly and disassembly using chemically defined components. *Proc. Natl Acad. Sci. USA* **100**, 8253–8257 (2003).
- Antonny, B., Huber, I., Paris, S., Chabre, M. & Cassel, D. Activation of ADP-ribosylation factor 1 GTPase-activating protein by phosphatidylcholine-derived diacylglycerols. *J. Biol. Chem.* **272**, 30848–30851 (1997).
- Stammes, M., Schiavo, G., Stenbeck, G., Sollner, T. H. & Rothman, J. E. ADP-ribosylation factor and phosphatidic acid levels in Golgi membranes during budding of coatamer-coated vesicles. *Proc. Natl Acad. Sci. USA* **95**, 13676–13680 (1998).
- Yanagisawa, L. L. *et al.* Activity of specific lipid-regulated ADP ribosylation factor-GTPase-activating proteins is required for Sec14p-dependent Golgi secretory function in yeast. *Mol. Biol. Cell* **13**, 2193–2206 (2002).
- Mayer, L. D., Hope, M. J. & Cullis, P. R. Vesicles of variable sizes produced by a rapid extrusion procedure. *Biochim. Biophys. Acta* **858**, 161–168 (1986).
- Brown, M. T. *et al.* ASAP1, a phospholipid-dependent arf GTPase-activating protein that associates with and is phosphorylated by Src. *Mol. Cell. Biol.* **18**, 7038–7051 (1998).
- Randazzo, P. A. *et al.* The Arf GTPase-activating protein ASAP1 regulates the actin cytoskeleton. *Proc. Natl Acad. Sci. USA* **97**, 4011–4016 (2000).
- Spang, A., Matsuoka, K., Hamamoto, S., Schekman, R. & Orci, L. Coatamer, Arf1p, and nucleotide are required to bud coat protein complex I-coated vesicles from large synthetic liposomes. *Proc. Natl Acad. Sci. USA* **95**, 11199–11204 (1998).
- Bremser, M. *et al.* Coupling of coat assembly and vesicle budding to packaging of putative cargo receptors. *Cell* **96**, 495–506 (1999).
- Antonny, B., Madden, D., Hamamoto, S., Orci, L. & Schekman, R. Dynamics of the COPII coat with GTP and stable analogues. *Nature Cell Biol.* **3**, 531–537 (2001).
- Goldberg, J. Decoding of sorting signals by coatamer through a GTPase switch in the COPI coat complex. *Cell* **100**, 671–679 (2000).
- Chabre, M. Aluminumfluoride and berylliofluoride complexes: new phosphate analogs in enzymology. *Trends Biochem. Sci.* **15**, 6–10 (1990).
- Scheffzek, K., Ahmadian, M. R. & Wittinghofer, A. GTPase-activating proteins: helping hands to complement an active site. *Trends Biochem. Sci.* **23**, 257–262 (1998).
- Finazzi, D., Cassel, D., Donaldson, J. G. & Klausner, R. D. Aluminum fluoride acts on the reversibility of ARF1-dependent coat protein binding to Golgi membranes. *J. Biol. Chem.* **269**, 13325–13330 (1994).
- Zhu, Y., Traub, L. M. & Kornfeld, S. ADP-ribosylation factor 1 transiently activates high-affinity adaptor protein complex AP-1 binding sites on Golgi membranes. *Mol. Biol. Cell* **9**, 1323–1337 (1998).
- Antonny, B., Gounon, P., Schekman, R. & Orci, L. Self-assembly of minimal COPII cages. *EMBO Rep.* **4**, 419–424 (2003).
- Presley, J. F. *et al.* Dissection of COPI and Arf1 dynamics *in vivo* and role in Golgi membrane transport. *Nature* **417**, 187–193 (2002).
- Yang, J. S. *et al.* ARFGAP1 promotes the formation of COPI vesicles, suggesting function as a component of the coat. *J. Cell Biol.* **159**, 69–78 (2002).
- Franco, M., Chardin, P., Chabre, M. & Paris, S. Myristoylation of ADP-ribosylation factor 1 facilitates nucleotide exchange at physiological Mg²⁺ levels. *J. Biol. Chem.* **270**, 1337–1341 (1995).
- Nickel, W. & Wieland, F. T. Receptor-dependent formation of COPI-coated vesicles from chemically defined donor liposomes. *Methods Enzymol.* **329**, 388–404 (2001).
- Matsuoka, K. & Schekman, R. The use of liposomes to study COPII- and COPI-coated vesicle formation and membrane protein sorting. *Methods* **20**, 417–428 (2000).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank P. Randazzo and R. Premont for proteins and constructs, J. Moelleken for advice on coatamer purification and lipopeptide synthesis and M. Chabre for discussions. This work was supported by the CNRS (ACI 'Jeune Chercheur') and the EMBO Young Investigator Programme.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to B.A. (antonny@ipmc.cnrs.fr).

Targeted degradation of TOC1 by ZTL modulates circadian function in *Arabidopsis thaliana*

Paloma Más¹, Woe-Yeon Kim², David E. Somers² & Steve A. Kay¹

¹Department of Cell Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037, USA

²Department of Plant Biology, Plant Biotechnology Center, Ohio State University, Columbus, Ohio 43210, USA

The underlying mechanism of circadian rhythmicity appears to be conserved among organisms, and is based on negative transcriptional feedback loops forming a cellular oscillator (or 'clock')^{1,2}. Circadian changes in protein stability, phosphorylation and subcellular localization also contribute to the generation and maintenance of this clock^{1,2}. In plants, several genes have been shown to be closely associated with the circadian system^{3,4}. However, the molecular mechanisms proposed to regulate the plant clock are mostly based on regulation at the transcriptional level^{3,4}. Here we provide genetic and molecular evidence for a role of ZEITLUPE (ZTL)^{5–7} in the targeted degradation of TIMING OF CAB EXPRESSION 1 (TOC1)^{8,9} in *Arabidopsis thaliana* (thale cress). The physical interaction of TOC1 with ZTL is abolished by the *ztl-1* mutation, resulting in constitutive levels of TOC1 protein expression. The dark-dependent degradation of TOC1 protein requires functional ZTL, and is prevented by inhibiting the proteasome pathway. Our results show that the TOC1–ZTL interaction is important in the control of TOC1 protein stability, and is probably responsible for the regulation of circadian period by the clock.

TOC1 is involved in a transcriptional feedback loop¹⁰ to control its own expression and that of CCA1 and LHY, two Myb DNA binding proteins associated with the clock^{11,12}. TOC1 has also been identified as a link between environmental information and clock outputs, with an essential function in constant darkness and in the integration of red light signals to the clock¹³. Two members of the ZTL protein family (ZTL^{5–7} and LKP2¹⁴) have also been implicated in the circadian clockwork. Sequence similarity of their LOV (light, oxygen and voltage) domains with those of WHITE COLLAR-1 (WC-1) and the phototropins suggests they may bind flavin and function as light sensors in *Arabidopsis*^{15,16}. The presence of a Kelch domain and an F-box motif implies their participation in E3 ubiquitin ligase SCF complexes^{17,18}. However, no functional evidence has been reported for any member of the ZTL family in targeting specific substrates for degradation.

Our previous studies have shown that a TOC1 RNAi transgene reduces TOC1 messenger RNA levels and shortens the free-running period of circadian gene expression in continuous white light¹³, whereas increased TOC1 gene dosage (using a TOC1 minigene (TMG) that expresses TOC1 under its endogenous promoter) delays the pace of the clock¹³. Conversely, lengthening of the free-running period was reported for loss-of-function and null *ztl* mutants^{5,6}. We therefore explored the inverse phenotypic correlation between TOC1 and ZTL expression by conducting a genetic study in which *ztl-1* mutant plants were transformed with either the TOC1 RNAi or the TMG constructs¹³. The *ztl-1* mutant plants exhibit circadian and photomorphogenic phenotypes similar to the null *ztl-3/ado* mutants⁶ (data not shown). Under constant white light (LL), the *ztl-1/TOC1* RNAi plants displayed circadian phenotypes similar to those observed in TOC1 RNAi plants¹³, with period lengths 3–4 h shorter than wild type (Fig. 1a, b). These results indicate that the *ztl-1* phenotype is only apparent in the presence of a functional TOC1, and is consistent with TOC1 acting downstream

of ZTL. In *ztl-1/TMG*, *CAB2::LUC* expression displayed circadian periods longer and with higher relative amplitudes error (that is, weaker rhythmicity) than those observed in either the single *ztl-1* mutant or TMG plants (Fig. 1c, d). Similar results were obtained with *ztl-3/TMG* plants; *CAB2::LUC* expression was clearly affected, with very long circadian periods and weak rhythms (Supplementary Fig. 1a, b). These results indicate that the long-period phenotype of TMG plants was enhanced by the *ztl* mutations. These findings are consistent with the notion that ZTL might target TOC1 for degradation.

We next examined the possible interaction between TOC1 and ZTL using the yeast two-hybrid system. Our results showed high lacZ activity in yeast colonies co-expressing TOC1 and ZTL, indicating a strong interaction between both proteins. Weaker interactions were detected with the other members of the ZTL

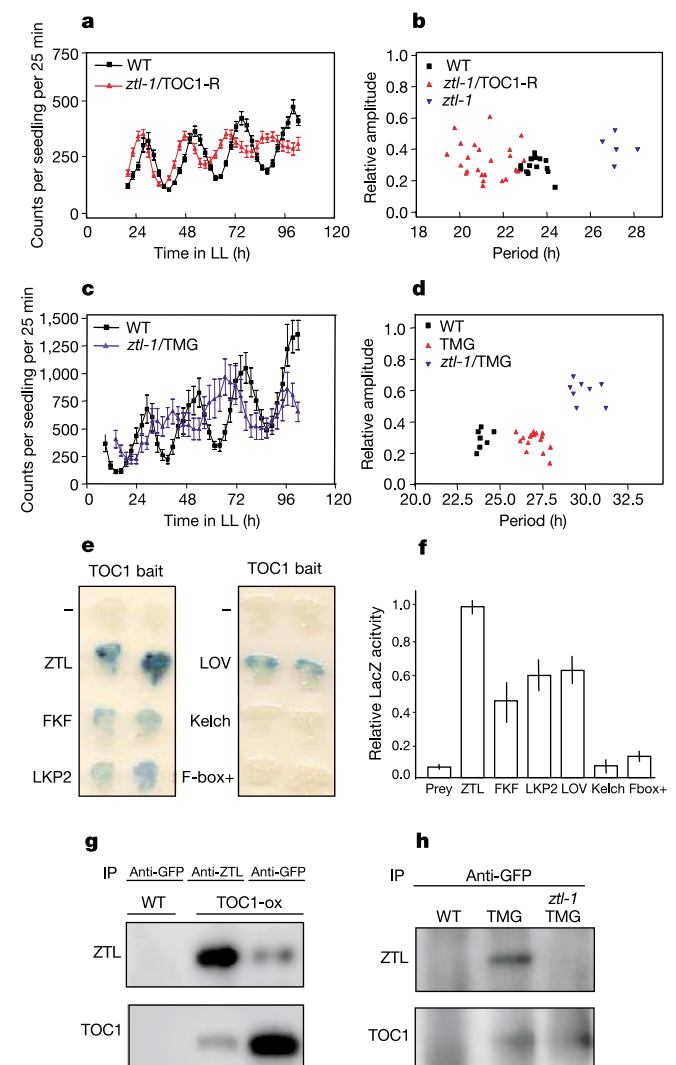


Figure 1 Genetic and molecular interaction between TOC1 and ZTL. **a, c**, Bioluminescence analysis of *CAB2::LUC* expression in LL. **b, d**, Period estimates of *CAB2::LUC* expression analysed as described^{21,22}. Seedlings were grown in LD cycles before transferring to LL. The experiments were repeated three times. WT, wild-type; TOC1-R, TOC1 RNAi; TMG, TOC1 minigene. **e**, TOC1 interaction in yeast with the ZTL family of proteins, and with the indicated domains of ZTL. —, empty prey vector; F-box +, F-box and Kelch domains. **f**, β -galactosidase activities in yeast. Values are the average of three colonies and the error bars represent standard deviations. **g, h**, Western blot analysis of immunoprecipitation with anti-GFP antibody and detection with antibody to ZTL²³ (see Methods). Plants were grown in LD cycles and processed at ZT-14.

family, FKF1 and LKP2 (Fig. 1e, f). Using different domains of ZTL, we found that the LOV domain was sufficient for the interaction with TOC1, whereas the Kelch or the F box + Kelch domains failed to interact in yeast (Fig. 1e, f). PAS domains are conserved among circadian systems, and have been identified as protein–protein interaction domains in a number of sensing proteins¹⁹. The PAS domains mediate interaction between WC-1 and WC-2, and such interactions are essential for circadian function in *Neurospora*²⁰. To verify the physical interaction between TOC1 and ZTL *in vivo*, we performed co-immunoprecipitation assays with TOC1–yellow fluorescent protein (YFP) overexpressing plants¹³ (TOC1-ox) which express TOC1–YFP under the strong 35S promoter. The YFP was used as a tag owing to the lack of a specific TOC1 antibody. Immunoprecipitation with anti-GFP antibody and subsequent detection of ZTL (Fig. 1g) showed a protein of molecular mass approximately 66 kDa coincident with the size of the protein immunoprecipitated with the ZTL antibody (lane anti-ZTL, Fig. 1g). Detection using the GFP antibody revealed a protein band with a molecular mass of 110 kDa that coincided with the predicted molecular size of the TOC1–YFP fusion protein. The physical interaction between TOC1 and ZTL was also observed in TMG plants (Fig. 1h). The TMG plants express TOC1–YFP under the TOC1 native promoter. Interestingly, we were unable to immunoprecipitate ZTL when *ztl-1*/TMG plants were assayed (Fig. 1h), although mutant *ztl-1* protein is expressed near wild-type levels (data not shown). These results indicate that the physical interaction between TOC1 and ZTL is abolished by the *ztl-1* mutation. Our yeast two-hybrid experiments show that the LOV domain is involved in the interaction of ZTL with TOC1. The *ztl-1* mutation in the Kelch domain⁵ may affect the conserved tertiary structure of the protein and disrupt the interaction with TOC1. The

lack of interaction between ZTL and TOC1 correlates with, and most probably accounts for, the circadian phenotypes observed in *ztl-1* and in *ztl-1*/TMG plants.

The genetic and molecular interaction between TOC1 and ZTL suggested that TOC1 protein levels might be regulated by ZTL. Time-course analysis in TMG seedlings grown in 12 h:12 h light:dark (LD) cycles revealed a rhythmically expressed TOC1 protein that closely followed the rhythmic pattern of mRNA expression (Fig. 2a, b). In *ztl-1*/TMG plants, the rhythmicity of TOC1 protein abundance was disrupted, with high levels at all time points assayed (compare Fig. 2a and c). Notably, the effects of the *ztl-1* mutation were specific for TOC1 protein and did not alter the overall rhythm of TOC1 mRNA expression (Fig. 2d). Together, these results support the notion that a post-translational mechanism, involving a functional ZTL, regulates the changes in TOC1 protein levels.

TMG seedlings displayed a long-period phenotype of *CAB2::LUC* expression in LL (Fig. 1d). Compared to LD cycles, TOC1 rhythmic accumulation in LL was delayed, with reduced amplitude and higher protein levels at the expected trough in the subjective day (Fig. 3a, b). In *ztl-1*/TMG (Fig. 3c) and *ztl-3*/TMG (Supplementary Fig. 1c), TOC1 protein expression was consistently high, with no rhythmic variations throughout the sampling interval. Interestingly, analysis of different transgenic lines revealed that the severity of the long-period circadian phenotypes correlated with higher levels of TOC1 protein (Fig. 3d). In *ztl-1*/TMG lines expressing very high levels of TOC1 protein, *CAB2::LUC* was arrhythmic and no circadian period (in the range of 15 to 45 h) was obtained after fast-Fourier transform-nonlinear least squares (FFT-NLLS) analysis^{21,22}. These results indicate that light/dark transitions and a functional ZTL are important in the control of TOC1 protein abundance.

We next made use of a cell-free assay to examine TOC1 *in vitro*

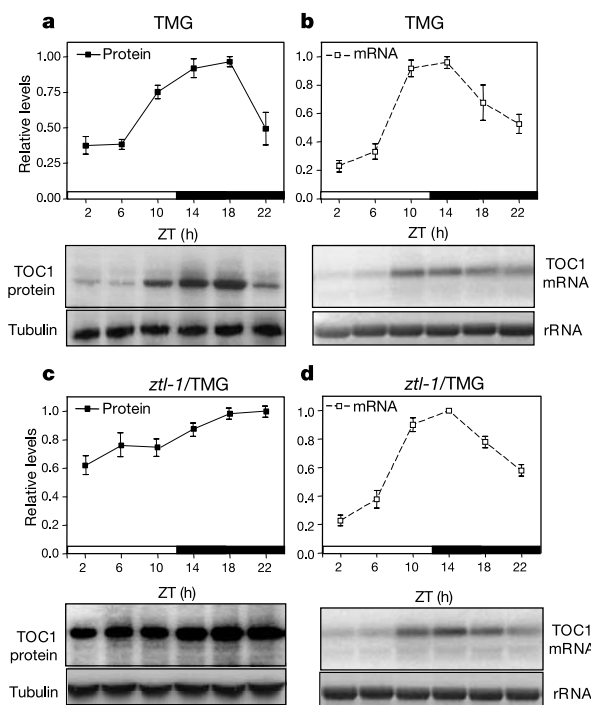


Figure 2 TOC1 protein and RNA expression in LD cycles. **a, c**, Immunodetection of TOC1 protein using anti-GFP antibody (see Methods) in TMG (**a**) and *ztl-1*/TMG (**c**) seedlings. Samples were collected at the indicated ZT (ZT-0 is light-on). TOC1 immunoreactivity was quantified relative to the maximum value for each blot after normalization to the tubulin. **b, d**, TOC1 mRNA in TMG (**b**) and *ztl-1*/TMG (**d**) plants analysed by northern blot. Samples were processed as described in Methods. rRNA was used as control. Data are mean of at least two independent experiments. White and dark bars represent light and dark period, respectively.

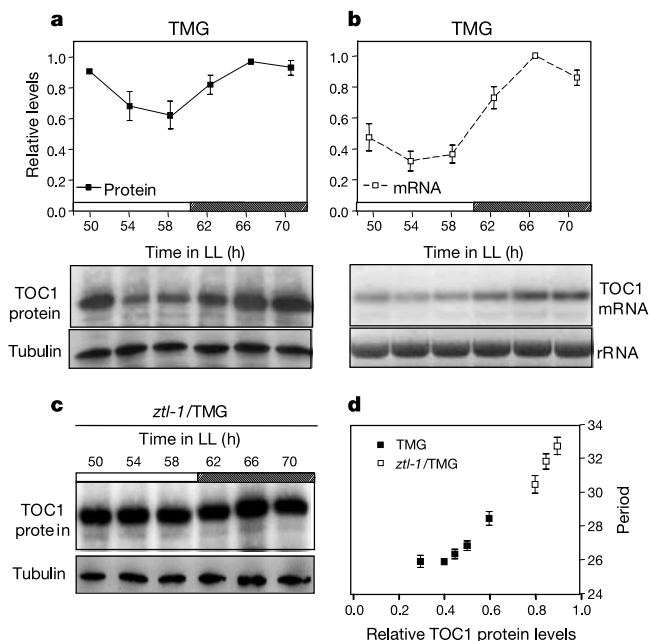


Figure 3 TOC1 RNA and protein expression in LL. **a, c**, Immunodetection of TOC1 protein using anti-GFP antibody (see Methods) in TMG (**a**) and *ztl-1*/TMG (**c**) seedlings grown in LD cycles before transferring to LL. Samples were collected at the indicated ZT (ZT-0 is light-on) and processed at the indicated time after the last LD light-off. **b**, TOC1 mRNA in TMG seedlings analysed by northern blot. Samples were processed as described in Methods. rRNA was used as control. Data are mean of at least two independent experiments. White bars represent light period; hatched bars indicate subjective night. **d**, Correlation between circadian period of *CAB2::LUC* expression and TOC1 protein levels in TMG and *ztl-1*/TMG lines. Seedlings were grown in LD cycles before transferring to LL. Period estimates were analysed as described^{21,22}.

degradation at the time of maximum (zeitgeber time, ZT-19) and minimum (ZT-7) TOC1 protein expression. In TMG extracts assayed at ZT-19, TOC1 protein levels were significantly reduced within 30 min (Fig. 4a). This degradation was prevented by using proteasome-specific inhibitors (Fig. 4a and Supplementary Fig. 2). The protease inhibitors aprotinin and leupeptin had little effect in preventing TOC1 degradation (Supplementary Fig. 2). Interestingly, in *ztl-1*/TMG extracts, TOC1 protein levels showed little change with no evident effect on protein abundance after adding proteasome (MG115 and MG132) or protease (aprotinin, leupeptin) inhibitors (Fig. 4a and Supplementary Fig. 2). When samples were processed at ZT-7, no degradation was observed in either TMG or *ztl-1*/TMG plants (Fig. 4b). The use of cycloheximide, a protein-synthesis inhibitor, revealed that the rate of TOC1 protein decrease was significantly higher at ZT-19 than at ZT-7 (data not shown), confirming that protein degradation rather than RNA translation is primarily responsible for the differential abundance of TOC1 protein. These results suggest that the rapid decrease in TOC1 protein levels from ZT-19 (Fig. 2a) occurs by degradation via the proteasome pathway and requires a functional ZTL protein.

To investigate a possible light- or dark-dependent TOC1 degradation, protein levels were analysed on transfer to constant darkness (DD, Fig. 5a) or constant light (LL, Fig. 5b). Extended darkness at ZT-24 into the first subjective day caused a clear reduction in TOC1 levels as compared to those observed in LD-grown seedlings (Fig. 5a). In contrast, when the light period was extended into the first subjective night, high levels of the TOC1 protein were maintained at all circadian time (CT) points analysed (Fig. 5b). Upon transfer to LL (not shown) or DD conditions (Fig. 5c, d), no evident changes in TOC1 protein levels were observed in *ztl-1*/TMG or *ztl-3*/TMG seedlings as compared with corresponding levels in LD cycles. These findings indicate that independent of its circadian regulation, TOC1 is preferentially degraded in the dark and ZTL plays a major role in that degradation. ZTL protein levels cycle with peak expression at approximately ZT-10 to ZT-12²³. The similar phase of TOC1 and ZTL expression suggests that additional factors might be involved in the regulation of TOC1. It is possible that light signals modulate changes in the binding affinity of the LOV domain and/or in the target degradation activity of ZTL. In LD, the light period appears to delay for some hours into the dark period the ability of ZTL to effect TOC1 degradation. In LL, only the rapid degradation mediated by ZTL is affected, so that TOC1 protein oscillates albeit with a lower amplitude. This situation is analogous to that of TIM in *Drosophila*, where TIM is degraded in the light by a CRY-dependent mechanism, but the protein still cycles in DD although with reduced

amplitude². We propose that both the circadian control of TOC1 mRNA expression and the ZTL-dependent TOC1 protein degradation act together to establish the rhythmic changes in TOC1 levels.

The alternation of day and night cycles has shaped the circadian clockwork throughout evolution. In plants, transmission of light signals to the core of the oscillator relies on the overlapping activities of different photoreceptors^{24,25}. The mechanisms and signalling pathways from this photoreception to the circadian outputs are just beginning to be understood. TOC1 expression is controlled by both transcriptional and post-translational mechanisms. In the former, the morning-expressed CCA1/LHY proteins bind to the evening-element sequence present in the TOC1 promoter, repressing its expression at dawn¹⁰. Decreasing CCA1/LHY levels relieve that repression, resulting in rising levels of TOC1 mRNA during the day. A second regulation at the post-translational level occurs with the activation of ZTL that targets TOC1 for degradation. Both levels of regulation ensure a narrow peak of TOC1 expression that might be critical in maintaining the robustness and stability of the clock. Evidence showing the involvement of the proteasome pathway at the core of the oscillator in the mammalian system²⁶ and in *Drosophila*^{27,28} suggests that tightly regulated protein degradation by the proteasome might be a

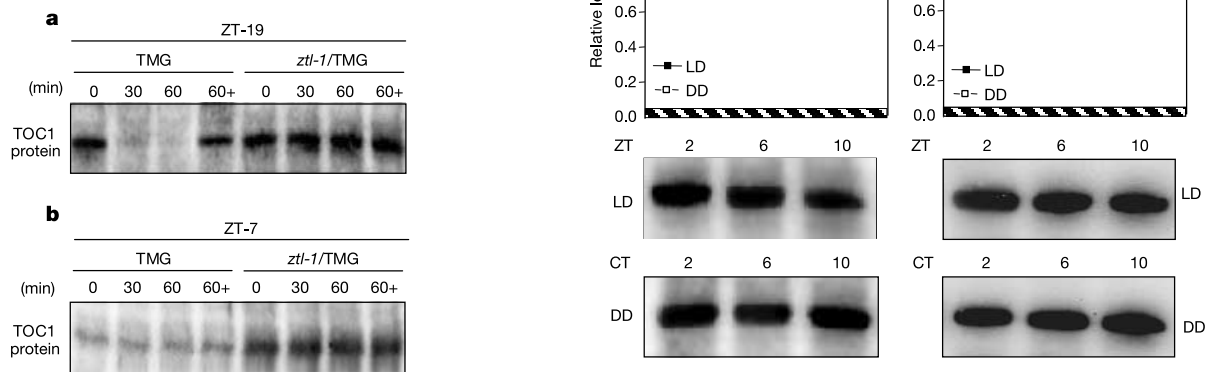


Figure 4 Cell-free TOC1 protein degradation assay. Immunodetection of TOC1 protein in TMG and *ztl-1*/TMG seedlings grown in LD cycles. Samples were collected at ZT-19 (a) and ZT-7 (b). Protein extracts were incubated in an *in vitro* degradation buffer (see Methods) with or without inhibitors for the indicated time (min). Approximately 80 μ g (TMG) and 40 μ g (*ztl-1*/TMG) of total proteins were loaded onto gels. 60+ indicates extracts treated with 50 μ M MG132.

Figure 5 Dark-dependent TOC1 protein degradation. Comparison of TOC1 protein in seedlings grown in LD cycles and on transfer to DD (a, c, d) or LL (b). Protein levels were analysed at ZT-2, ZT-6, ZT-10, ZT-14 or their corresponding circadian times (CT). Approximately 80 μ g of total protein was loaded in TMG extracts (a and b) whereas 50 μ g was loaded in *ztl-1*/TMG (c) and *ztl-3*/TMG (d). White and dark bars represent light and dark period, respectively. Hatched bar in a, c and d indicates subjective day; hatched bar in b indicates subjective night.

conserved aspect in the regulation of the eukaryotic circadian clocks. □

Methods

Plant growth conditions and bioluminescence analysis

Arabidopsis plants transformed by *Agrobacterium*-mediated DNA transfer were selected in MS-agar, 3% sucrose plates. Seedlings were grown under 12 h:12 h light:dark (LD) cycles with 60 $\mu\text{mol m}^{-2} \text{s}^{-1}$ of cool white fluorescent light at 22 °C. Analysis of bioluminescence rhythms under continuous white light was performed as previously described^{29,30}. FFT-NLLS analysis was used to estimate period lengths of individual traces^{21,22}. *TOC1* minigene (TMG–YFP) plants express the *TOC1* genomic sequence (from 2,354 base pairs (bp) upstream of the ATG to 410 bp downstream of the stop codon) fused to the YFP. The *TOC1* RNAi and *TOC1-ox* constructs were previously described¹³.

Yeast two-hybrid assays

The coding region of *TOC1* complementary DNA was cloned into the pBMT116 vector (Clontech) as a translational fusion to the LexA DNA binding protein and used as a bait for the yeast two-hybrid studies. ZTL full-length and the domains LOV/PAS (residues 1–155), Kelch (residues 290–609) and F-box + Kelch (residues 200–609) were cloned into the pACT2 vector (Clontech) and used as prey. The yeast transformation and enzymatic activity on plates and liquid media was performed following manufacturer recommendations (Clontech Yeast Protocols Handbook).

Coimmunoprecipitation experiments

Wild-type, *TOC1-ox*, TMG and *ztl-1*/TMG plants were grown in 12 h:12 h LD cycles for 15 days and processed for co-immunoprecipitation at ZT-14. *TOC1*–YFP immunoprecipitation was performed at 4 °C for 4 h using anti-GFP antibody (Molecular Probes) in a buffer containing 50 mM Tris–HCl pH 7.5, 150 mM NaCl; 0.5% NP-40, 1 mM EDTA, 3 mM dithiothreitol (DTT), 1 mM phenylmethylsulfonyl fluoride (PMSF), 5 $\mu\text{g ml}^{-1}$ leupeptin, 1 $\mu\text{g ml}^{-1}$ aprotinin, 5 $\mu\text{g ml}^{-1}$ antipain, 1 $\mu\text{g ml}^{-1}$ pepstatin, 5 $\mu\text{g ml}^{-1}$ chymostatin and 50 μM MG132. Protein A/agarose beads were used to precipitate the immunoprotein complexes. ZTL detection was performed using anti-ZTL antibody²³.

Western and northern blots

Proteins were extracted from 0.2 g of tissue and ground in 0.5 ml of grinding buffer (50 mM HEPES–KOH pH 7.5, 100 mM KCl, 10% glycerol, 0.1% Triton X-100, 2 mM DTT, 5 mM EDTA, 0.5% polyvinylpyrrolidone (PVPP), 50 μM MG132 and Protease Inhibitor Cocktail (Roche)). Protein concentration was determined using the Bradford method (protein assay, Bio-Rad) and approximately 70–80 μg of total protein was loaded per lane. Homogeneous protein transference to nitrocellulose membranes was confirmed by red Ponceau staining. Anti-GFP (Molecular Probes) and anti-ZTL antibodies²³ were used to detect *TOC1*–YFP and ZTL protein, respectively. As control, membranes were incubated with anti-tubulin antibody (Sigma). For northern blot analysis, RNA from 12-day-old seedlings was isolated using the RNeasy kit (Qiagen) and separated on 1.2% agarose/formaldehyde gels as previously described¹³. Northern blots were performed using radioactively labelled full-length YFP DNA fragment as probe. Quantification of the blots and analysis of the images was performed on a Phosphorimager and using ImageQuant software (Molecular Dynamics) and Scion Image software (Scion Corporation).

Cell-free degradation assays

Seedlings were ground in liquid nitrogen, resuspended in buffer (50 mM Tris–HCl pH 7.5, 100 mM NaCl, 10 mM MgCl_2 , 5 mM DTT and 5 mM ATP) and clarified by centrifugation. Equal amounts of extracts were transferred to individual tubes and incubated at room temperature with or without inhibitors as indicated. Reactions were stopped by adding protein gel-loading buffer.

Received 17 July; accepted 17 October 2003; doi:10.1038/nature02163.

- Harmer, S. L., Panda, S. & Kay, S. A. Molecular bases of circadian rhythms. *Annu. Rev. Cell Dev. Biol.* **17**, 215–253 (2001).
- Young, M. W. & Kay, S. A. Time zones: A comparative genetics of circadian clocks. *Nature Rev. Genet.* **2**, 702–715 (2001).
- McClung, C. R. Circadian rhythms in plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **52**, 139–162 (2001).
- Yanovsky, M. J. & Kay, S. A. Signaling networks in the plant circadian system. *Curr. Opin. Plant Biol.* **4**, 429–435 (2001).
- Somers, D. E., Schultz, T. F., Milnamow, M. & Kay, S. A. ZEITLUPE encodes a novel clock-associated PAS protein from *Arabidopsis*. *Cell* **101**, 319–329 (2000).
- Jarillo, J. A. *et al.* An *Arabidopsis* circadian clock component interacts with both CRY1 and phyB. *Nature* **410**, 487–490 (2001).
- Kiyosue, T. & Wada, M. LKP1 (LOV kelch protein 1): a factor involved in the regulation of flowering time in *Arabidopsis*. *Plant J.* **23**, 807–815 (2000).
- Strayer, C. A. *et al.* Cloning of the *Arabidopsis* clock gene *TOC1*, an autoregulatory response regulator homolog. *Science* **289**, 768–771 (2000).
- Makino, S., Matsushika, A., Kojima, M., Yamashino, T. & Mizuno, T. The APRR1/TOC1 quintet implicated in circadian rhythms of *Arabidopsis thaliana*: Characterization with APRR1-overexpressing plants. *Plant Cell Physiol.* **43**, 58–69 (2002).
- Alabadi, D. *et al.* Reciprocal regulation between *TOC1* and *LHY/CCA1* within the *Arabidopsis* circadian clock. *Science* **293**, 880–883 (2001).
- Wang, Z. Y. & Tobin, E. M. Constitutive expression of the CIRCADIAN CLOCK ASSOCIATED 1 (CCA1) gene disrupts circadian rhythms and suppresses its own expression. *Cell* **93**, 1207–1217 (1998).

- Schaffer, R. *et al.* The late elongated hypocotyl mutation of *Arabidopsis* disrupts circadian rhythms and the photoperiodic control of flowering. *Cell* **93**, 1219–1229 (1998).
- Más, P., Alabadi, D., Yanovsky, M. J., Oyama, T. & Kay, S. A. Dual role of *TOC1* in the control of circadian and photomorphogenic responses in *Arabidopsis*. *Plant Cell* **15**, 223–236 (2003).
- Schultz, T. F., Kiyosue, T., Yanovsky, M., Wada, M. & Kay, S. A. A role for LKP2 in the circadian clock of *Arabidopsis*. *Plant Cell* **13**, 2659–2670 (2001).
- Cheng, P., He, Q., Yang, Y., Wang, L. & Liu, Y. Functional conservation of light, oxygen, or voltage domains in light sensing. *Proc. Natl Acad. Sci. USA* **100**, 5938–5943 (2003).
- Briggs, W. R. & Christie, J. M. Phototropins 1 and 2: Versatile plant blue-light receptors. *Trends Plant Sci.* **7**, 201–210 (2002).
- Deshaires, R. J. SCF and cullin/RING H2-based ubiquitin ligases. *Annu. Rev. Cell Dev. Biol.* **15**, 435–467 (1999).
- Risseuw, E. P. *et al.* Protein interaction analysis of SCF ubiquitin E3 ligase subunits from *Arabidopsis*. *Plant J.* **34**, 753–767 (2003).
- Taylor, B. L. & Zhulin, I. B. PAS Domains: Internal sensors of oxygen, redox potential, and light. *Microbiol. Mol. Biol. Rev.* **63**, 479–506 (1999).
- Froehlich, A. C., Liu, Y., Loros, J. J. & Dunlap, J. C. White Collar-1, a circadian blue light photoreceptor, binding to the frequency promoter. *Science* **297**, 815–819 (2002).
- Plautz, J. D. *et al.* Quantitative analysis of *Drosophila* period gene transcription in living animals. *J. Biol. Rhythms* **12**, 204–217 (1997).
- Millar, A. J., Straume, M., Chory, J., Chua, N. H. & Kay, S. A. The regulation of circadian period by phototransduction pathways in *Arabidopsis*. *Science* **267**, 1163–1166 (1995).
- Kim, W.-Y., Geng, R. & Somers, D. E. Circadian phase-specific degradation of the F-box protein ZTL is mediated by the proteasome. *Proc. Natl Acad. Sci. USA* **100**, 4933–4938 (2003).
- Somers, D. E., Devlin, P. F. & Kay, S. A. Phytochromes and cryptochromes in the entrainment of the *Arabidopsis* circadian clock. *Science* **282**, 1488–1490 (1998).
- Más, P., Devlin, P. F., Panda, S. & Kay, S. A. Functional interaction of phytochrome B and cryptochrome 2. *Nature* **408**, 207–211 (2000).
- Yagita, K. *et al.* Nucleocytoplasmic shuttling and mCRY-dependent inhibition of ubiquitylation of the mPER2 clock protein. *EMBO J.* **21**, 1301–1314 (2002).
- Ko, H. W., Jiang, J. & Edery, I. Role for Slimb in the degradation of *Drosophila* Period protein phosphorylated by Doubletime. *Nature* **420**, 673–678 (2002).
- Grima, B. *et al.* The F-box protein slimb controls the levels of clock proteins period and timeless. *Nature* **420**, 178–182 (2002).
- Millar, A. J., Carre, I. A., Strayer, C. A., Chua, N. H. & Kay, S. A. Circadian clock mutants in *Arabidopsis* identified by luciferase imaging. *Science* **267**, 1161–1163 (1995).
- Somers, D. E., Webb, A. A. R., Pearson, M. & Kay, S. A. The short-period mutant *tocl-1*, alters circadian clock regulation of multiple outputs throughout development in *Arabidopsis thaliana*. *Development* **125**, 485–494 (1998).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank H. Tran, T. Imaizumi and S. Hazen for critical reading of the manuscript. This research was supported by the NIH (S.A.K.) and the NSF (D.E.S.).

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to S.A.K. (stevek@scripps.edu).

Proteomic characterization of the human centrosome by protein correlation profiling

Jens S. Andersen^{1*}, Christopher J. Wilkinson^{2*}, Thibault Mayor^{2†}, Peter Mortensen¹, Erich A. Nigg² & Matthias Mann¹

¹Center for Experimental BioInformatics (CEBI), Department of Biochemistry and Molecular Biology, University of Southern Denmark, Campusvej 55, DK-5230 Odense M, Denmark

²Department of Cell Biology, Max Planck Institute for Biochemistry, D-82152 Martinsried, Germany

* These authors contributed equally to this work

† Present address: Division of Biology MC 156-29, California Institute of Technology, 1200 East California Blvd, Pasadena, California 91125, USA

The centrosome is the major microtubule-organizing centre of animal cells and through its influence on the cytoskeleton is involved in cell shape, polarity and motility. It also has a crucial function in cell division because it determines the poles of the mitotic spindle that segregates duplicated chromosomes between

dividing cells¹⁻⁵. Despite the importance of this organelle to cell biology and more than 100 years of study, many aspects of its function remain enigmatic and its structure and composition are still largely unknown. We performed a mass-spectrometry-based proteomic analysis of human centrosomes in the interphase of the cell cycle by quantitatively profiling hundreds of proteins across several centrifugation fractions. True centrosomal proteins were revealed by both correlation with already known centrosomal proteins and *in vivo* localization. We identified and validated 23 novel components and identified 41 likely candidates as well as the vast majority of the known centrosomal proteins in a large background of nonspecific proteins. Protein correlation profiling permits the analysis of any multiprotein complex that can be enriched by fractionation but not purified to homogeneity.

Proteomics based on mass spectrometry⁶ has become a powerful method to determine the composition of complex protein mixtures. Nevertheless, the analysis of many cellular structures remains challenging because such structures are difficult to purify. Human centrosomes—tiny, single-copy organelles composed of a pair of cylindrical centrioles embedded in an electron-dense matrix of pericentriolar material—exemplify these impediments because they make up a small proportion of total cell protein, lack delineating membranes and often associate tightly with the nuclear membrane. In contrast, the cognate organelle in yeast, the spindle pole body, yielded to proteomic analysis several years ago⁷.

We purified the centrosome from the human lymphoblastic cell line KE-37 by using established protocols^{8,9}. The fractions from the final sucrose gradient were analysed by mass-spectrometry-based proteomics as described in Methods and Fig. 1. Sequence database searches resulted in the identification of more than 2,000 peptides representing more than 500 proteins in the peak centrosome fraction (Supplementary Tables A and B). The current literature lists about 60 proteins associated with the interphase (G1/S-phase) human centrosome. Of these, our mass spectrometric analysis detected 47. We also identified another 90 proteins that were previously characterized at the complementary DNA or genetic level only.

To determine which of the above proteins were genuine centrosomal components, we used two complementary approaches. First, we cloned 32 of the 90 previously uncharacterized proteins, selected primarily by the availability of cDNA clones. Each protein was expressed in U2OS cells from three different constructs providing a green fluorescent protein (GFP) tag at either the amino or carboxy terminus or a Myc epitope at the N terminus. (Immuno)fluorescence microscopy showed that 19 of the 32 candidate centrosomal proteins localized to the organelle (see Methods and Fig. 2). Unequivocal centrosome localization was seen at low expression levels. At higher levels, additional diffuse staining was seen throughout the cytoplasm for about half of these proteins. The others—all coiled-coil in structure—also formed aggregates, one of which invariably contained the centrosome. Two proteins (KIAA1633 and FLJ22445) in addition bound to microtubules. The localization studies were repeated on cells that had been treated with nocodazole. This showed that none of 15 proteins thus analysed required intact microtubules for centrosome localization (Fig. 2, and data not shown). Thus, by these criteria, the proteins characterized here are genuine centrosome components. The newly identified centrosomal proteins without prior name are designated here by the acronym Cep (for 'centrosomal protein'), followed by the approximate relative molecular mass (M_r ; Table 1).

As a second approach aimed at distinguishing centrosomal proteins from nonspecific proteins, we developed a method to track the abundance of peptides through five sucrose fractions of the centrosome preparation. Although it is difficult to establish the absolute amount of different peptides from the mass spectrometric signal, the relative amount of the same peptide can be determined

quite accurately in consecutive analyses. The peptides that were sequenced in the peak centrosomal fraction were therefore located in other fractions by elution time fitting (Fig. 1f and Methods). Peptide elution profiles for more than 1,600 peptides were verified by manual inspection of the raw data. Figure 3a shows that all of the peptides of the known centrosomal protein (AKAP350) follow the same fractionation profile; it also shows that the individual peptides are quantified within a band of about 30%, so the protein profile should be significantly more accurate than this. Contaminating proteins have a fractionation profile very different from that of centrosomal proteins. Also drawn in Fig. 3b is a 'consensus profile' defined as the average of the profiles of three known centrosomal proteins (pericentrin, γ -tubulin and centrin 2).

The differential subcellular localization of closely related protein isoforms can be distinguished by this approach as long as peptides corresponding to the differences in isoforms are sequenced. As can be seen in Fig. 3c, 14-3-3 ϵ but not 14-3-3 τ correlates with the consensus profile of Fig. 3b. This is in agreement with a previous observation in which 14-3-3 ϵ but not 14-3-3 τ was localized to the centrosome¹⁰.

To quantify the degree of correlation with centrosomal marker proteins, we defined a 'goodness of fit' for all other proteins calculated as the normalized square deviation from the centrosomal

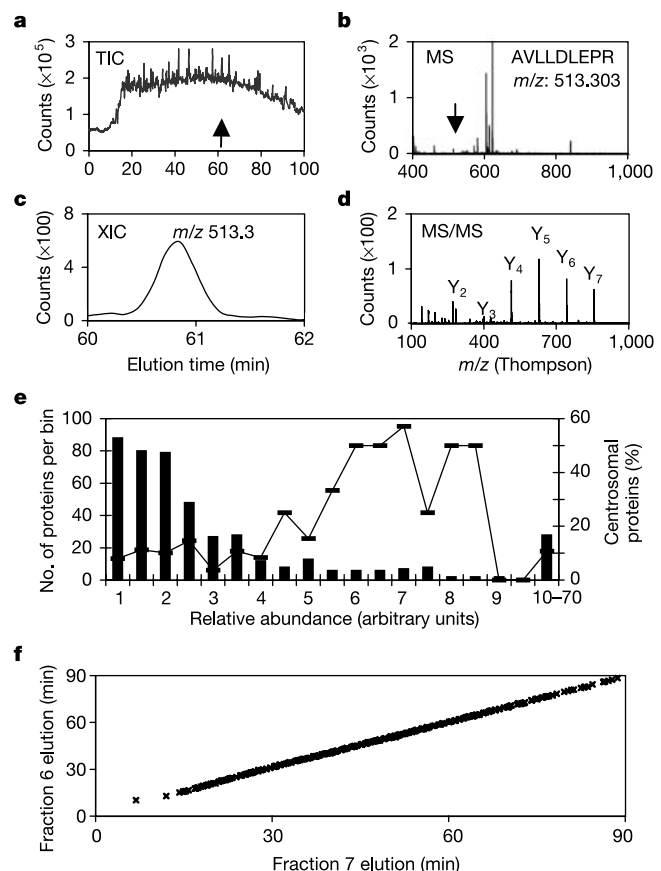


Figure 1 Proteomic analysis of the centrosome. **a**, LC MS/MS analysis of a centrifugation fraction. **b**, **d**, Mass spectrum (**b**) and fragmentation spectrum (**d**) of the γ -tubulin peptide. Predicted C-terminal fragments (Y) are marked. **c**, Ion current for the γ -tubulin peptide AVLLDLEPR at m/z 513.303 that eluted at the time shown in the chromatogram. The peak area is a measure of the abundance of the peptide. **e**, Relative abundances of proteins identified in the peak centrosomal fraction. Peptide signals were integrated as in **c** for the top three peptides to yield a rough estimate (to within a factor of four) of the copy number. The right-hand scale corresponds to the curve and indicates the fraction of centrosomal proteins per abundance bin. **f**, Fitting between LC MS/MS runs of different centrifugation fractions allowed accurate prediction of elution time.

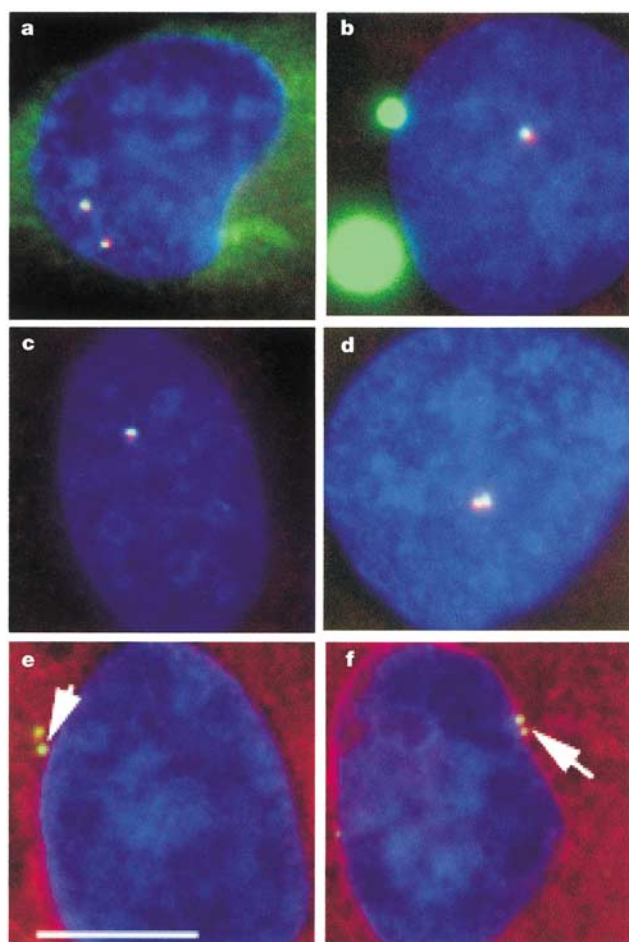


Figure 2 U2OS cells expressing a selection of novel centrosomal proteins as GFP fusion proteins. γ -Tubulin (centrosome marker, **a–d**) or α -tubulin (microtubule marker, **e, f**) was stained with Texas red and DNA with 4,6-diamidino-2-phenylindole; yellow indicates coincidence of green and red signals. **a**, ALMS1; **b**, Cep135 (KIAA0635); **c**, **e**, Cep152 (KIAA0912); **d**, **f**, Cep63 (FLJ13386). Samples in **e** and **f** were treated with nocodazole. Scale bar, 10 μ m.

consensus profile shown in Fig. 3b. In Fig. 3d this ' χ^2 value' is plotted for centrosomal proteins and a group of typical contaminating proteins, namely translation/initiation factors and ribosomal proteins. The two groups clearly separate in their χ^2 values, which therefore permits exclusion of the contaminating proteins on the basis of their fractionation behaviour. As a small-scale negative control of both the protein correlation profiling (PCP) and microscopy experiments, we tagged phosphofructokinase, phosphoglycerate dehydrogenase and cyclophilin A, which were identified in the peak fraction but are not expected to be centrosomal. Indeed, all three had high χ^2 values (Supplementary Table A) and the tagged proteins did not localize to the centrosome (data not shown).

Known centrosomal proteins had low χ^2 values, their profiles fitting well with those of the marker proteins. The vast majority of proteins that would not be predicted to be centrosomal on the basis of their ascribed function had high χ^2 values. Examples include ribosomal, membrane or nuclear proteins, metabolic enzymes and components of the spliceosome.

For some proteins only a single peptide could be quantified, or quantification could not be performed in all five fractions. To increase the mass spectrometric sensitivity and resolution further, we performed additional experiments in which the proteins in each centrifugal fraction were first separated by one-dimensional gel electrophoresis and 15 slices of each lane were quantified separately.

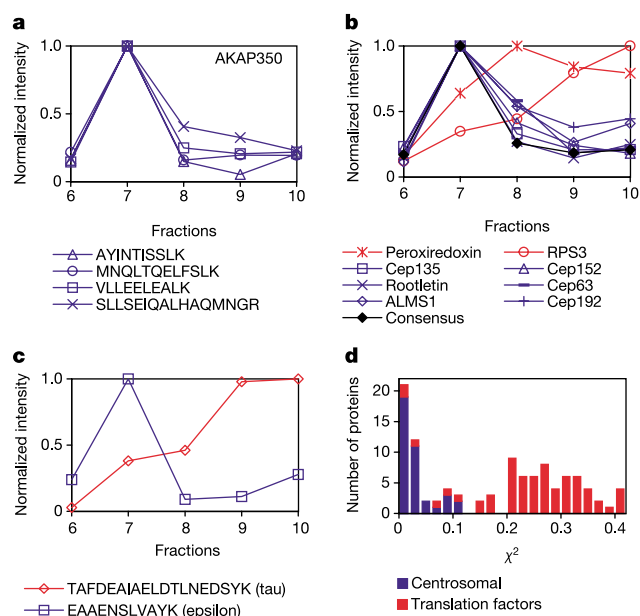


Figure 3 Protein correlation profiling of the human centrosome. **a**, Peptides identifying AKAP350 quantified through five centrifugation fractions. **b**, Protein elution profiles as an average of peptide profiles. Novel centrosomal proteins (blue) have a profile close to the 'consensus profile' (black) displayed by three known centrosomal proteins. Contaminating proteins (red) have different fractionation profiles. **c**, Isoform-specific detection of centrosomal proteins. 14-3-3 ϵ , but not τ , shows the same fractionation profile as the consensus profile in **b**. **d**, ' χ^2 ' values, calculated as the squared deviation from the consensus profile for centrosomal proteins (blue) and a group of contaminating proteins (red).

This experiment validated the initial results and provided protein correlation profiles for a large set of proteins. In another set of proteins, namely α - and β -tubulin, the heat shock proteins/complexes Hsp90 and TCP^{11,12}, a large cytoplasmic pool resulted in an 'offset' to the protein profile and to a high χ^2 value. Collecting and analysing more centrifugation fractions would furthermore increase the resolution of the method in cases in which protein complexes partly overlap.

To arrive at a list of candidate centrosomal components worthy of future investigation, we extracted 41 proteins with complete quantification data, protein χ^2 values of less than 0.05 and low χ^2 value for at least two peptides (Supplementary Table C).

In total, through a combination of mass spectrometric analysis and *in vivo* localization we have identified at least 70 genuine centrosome proteins as well as a group of likely candidates. Of the 23 new components, 19 were validated here by (immuno)localization and four were described independently by other groups in the course of this study (Table 1). We believe that we have obtained almost complete coverage of structural proteins, although we have not detected TACC2, protein 4.1R, Nlp and tubulins epsilon and delta^{13–16}. We have also identified at least two-thirds of the known regulatory proteins localized to the centrosome in interphase cells. Failure to detect the remaining proteins is probably due to low abundance, loss during purification or because these factors do not associate to a significant degree with the centrosome under the conditions of this experiment. Our study substantially increases the number of centrosomal proteins, and although this inventory is not yet complete it is striking that virtually all centrosomal proteins published since September 2002, when we initiated this screen, have already been detected in this work.

The novel proteins had M_r values in the range 13,000–461,000 (13K–461K), with an average of $\sim$ 130K. A structural analysis

Table 1 Proteins newly identified and shown to localize to human centrosomes

Accession no.	Protein name (clone)	Localization (GFP and Myc)	Localization (PCP)	Motifs, comments	M_r (K)	Relative abundance	Salt-stripped
NP_060567	Cep27* (FLJ10460)	Y	–	1 CC	27.1	69	P
NP_061188	Cep41 (FLJ22445)	Y	0.02	1 CC, rhodanese-like	41.3	173	S
Q9BVF9	Cep57 (KIAA0092)	Y	0.01	2 CC	57.4	684	P
Q9H8N0	Cep63 (FLJ13386)	Y	0.01	6 CC	63.4	494	P
NP_055962	Cep68 (KIAA0582)	Y	0.01		67.8	192	P
NP_077817	Cep70 (FLJ13036)	Y	0.04	2 CC, TPR	70.2	231	–
Q9P209	Cep72 (KIAA1519)	Y	0.03	LRR, LRRcap	71.8	86	–
NP_079175	Cep76 (FLJ12542)	Y	0.03		75.7	139	S
Q9H9N3	Cep78 (FLJ12643)	Y	–		78.3	161	–
Q9UPN4	Cep131 (KIAA1118)	Y	0.02	4 CC, troponin	130.6	435	P
NP_055460	Cep135 (KIAA0635)	Y	0.01	13 CC (suggested in ref. 25)	135	794	P
Q94986	Cep152 (KIAA0912)	Y	0.01	8 CC	151.5	746	P
NP_055771	Cep164 (KIAA1052)	Y	0.03	7 CC, WW†	164.0	127	–
NP_115518	Cep192 (KIAA1569 and FLJ00145)	Y	0.03	–	191.6	249	–
NP_060719	Cep215 (KIAA1633)	Y	0.05	9 CC homology to Rn CDK5 activator-binding protein, Rn Myomegalin	215	528	P
O15078	Cep290 (KIAA0373)	Y	0.02	9 CC homology to Hs CTCL tumour antigen	290	246	–
NP_055935	ALMS1	Y	0.01	Alstrom syndrome, disease linked	463.3	323	P
O75665	OFD-1	Y	0.01	4 CC, LisH (validated in ref. 22)	117.1	213	PS
O43805	NA-14	Y	0.09	1 CC (validated in ref. 26)	13.8	185	S
O60527	CCCAP	Y	–	5 CC (validated in ref. 27)	74.1	167	PS
O43303	CP110 (KIAA0419)	Y	0.01	2 CC (validated in ref. 28)	112.1	160	PS
NP_055490	Rootletin (KIAA0445)	Y	0.01	14 CC, 10 M-repeats (independently validated in ref. 29)	228	527	S
NP_008976	FOP	Y	–	LisH, chromosomal translocation, disease-linked	43.2	89	S

Novel centrosomal proteins were validated by fluorescence microscopy (GFP and Myc tag) and protein correlation profiling (PCP) where indicated. PCP is given in χ^2 values (see Methods). References are given for proteins independently verified elsewhere during this screen. Conserved motifs are abbreviated as follows: CC (coiled-coil), TPR (tetratricopeptide repeat), LisH (Lissencephaly type-1-like homology motif), WW (domain with two conserved Trp (W) residues), LRR (leucine-rich repeats), LRRcap (occurring C-terminal to leucine-rich repeats). Relative abundance is given in arbitrary units and is the average of extracted ion current of the three top peptides (J. Rappsilber and M.M., unpublished data). 'Salt-stripped' refers to protein found in the supernatant (S) or pellet (P) of the centrosomal fraction.

*Novel factors are termed Cep and a number, where Cep stands for centrosomal protein and is followed by the M_r calculated from the full-length sequence.

†The validation of Cep164 is tentative because only Myc-Cep164 was clearly centrosome-enriched.

revealed that a high proportion (75%) contained coiled-coil regions, a common feature of centrosomal proteins, and few had any other recognizable motifs or domains. This raises the intriguing question of how these many coiled-coil proteins cooperate to form a pericentriolar matrix.

The pericentriolar matrix can be subdivided into salt-insoluble and salt-soluble fractions¹⁷. We precipitated salt-extracted 'centrosomal scaffolds' and determined the relative amounts of the novel proteins in the supernatant and the pellet after centrifugation (see Methods). As expected from previous reports¹⁸, ODF-2 was found in the centrosomal scaffold with a ratio of greater than 10, whereas γ -tubulin and γ -tubulin ring-complex proteins were depleted by a factor of 5. This analysis was performed for all the proteins listed in Table 1; results are listed in Table 1 and Supplementary Tables A and C. About half of the proteins were tightly associated with the centrosomal scaffold by this assay.

Our results open a new perspective on several genes linked to human disease. Specifically, one protein identified here as a centrosomal component was found to be fused to fibroblast growth factor receptor in several cases of an atypical stem-cell myeloproliferative disorder with lymphoma (B or T cell), myeloid hyperplasia, and eosinophilia¹⁹. It will be interesting to determine whether this fusion protein affects centrosome function in the corresponding tumour cells. We have also found that the C-terminal half of the ALMS1 (Alström syndrome) protein, whose mutation causes obesity, type 2 diabetes and neurosensory degeneration^{20,21}, localizes to the centrosome, and the available evidence indicates that OFD1 (oral, facial, digital syndrome) is a genuine centrosomal protein²². Given that mutation in these genes causes developmental defects, it is tempting to speculate that a failure in the organization of the cytoskeleton by the centrosome in certain tissues could contribute to the aetiology of these diseases.

Previous studies on centrosome proteins were directed at particular functions and therefore identified only a limited number of components. In contrast, our comprehensive approach identified 23 new centrosome components and at least 41 additional likely

candidates. Future studies will study possible variations in the composition of the centrosome, namely during the cell cycle and in different cell types using the PCP method introduced here, possibly in combination with isotope-labelling techniques⁶, which can additionally reveal differences between two states. Moreover, in combination with new methods for assaying centrosome function (such as short interfering RNA) our data set the stage for a functional dissection of this hitherto mysterious organelle.

The PCP algorithm described here offers considerable promise for the study of other multiprotein complexes that are difficult to obtain in pure form, which encompasses most cellular structures. Our technique allows the requirement for complex purification to be relaxed significantly, avoiding the trade-off between rigorous purification and the inclusion of loosely associated but genuine components, and could be the basis of systematically defining the spatially resolved proteome of cell lines and tissues²³.

Note added in proof: KIAA0542, identified here as a candidate centrosomal protein in Supplementary Table C, has been confirmed and termed hSif1p³⁰. □

Methods

Isolation of centrosomes

Centrosomes were isolated from $(2-3) \times 10^9$ human lymphoblastic KE37 cells as described previously⁸. In brief, this method involves treating exponentially growing cells with nocodazole and cytochalasin D, followed by hypotonic lysis. Centrosomes are harvested by centrifugation onto a 50% sucrose cushion and further purified by centrifugation through a discontinuous (70%, 50% and 40%) sucrose gradient. Fractions of 0.5 ml were collected and analysed. For salt extraction (KI) experiments, the isolated centrosomes were first pelleted to remove sucrose and then treated with 200 μ l 2 M KI, 10 mM K-PIPES pH 7.2 for 2 h at 4 °C. The KI-treated centrosomes were pelleted at 16,000g for 10 min, and the pellet and the supernatant fractions were analysed. Protein mixtures were either digested directly or separated by one-dimensional polyacrylamide gel electrophoresis and excised in 15 equal slices.

Identification of proteins by NanoLC MS/MS

Isolated centrosomes were dissolved in 100 μ l 8 M urea, 100 mM Tris-HCl at pH 8. Proteins were reduced and alkylated, then digested with endoproteinase Lys-C for 4 h and with trypsin for 8 h after dilution with 400 μ l 50 mM NH_4HCO_3 pH 8.0. The resulting peptide mixture was desalted and concentrated, then pressure-loaded onto a pulled fused

silica capillary with an internal diameter of 100 μm and a tip opening of 8 μm (New Objective) filled with Vydac 218MSB3 3 μm reverse-phase material. Peptides were eluted directly into a quadrupole time-of-flight mass spectrometer (MDS Sciex) with a 90-min linear gradient of 5–60% buffer B (80% acetonitrile, 0.5% acetic acid, 0.02% heptafluorobutyric acid) from an HPLC system (Agilent). The nanoLC MS/MS experiments were repeated with precursor ions selected from restricted m/z ranges (m/z 350–550, 545–750 and 745–1,400) matched by pulsed extraction of the product ions. Combined peak lists were searched in the International Protein Index (IPI) database (<http://www.ebi.ac.uk/IPI/IPIhelp.html>) using the Mascot program (Matrix Science). Iterative calibration algorithms on the basis of identified peptides were used to achieve an average absolute mass accuracy of better than 20 p.p.m. in both precursor and product ions. Proteins identified with a combined peptide score of higher than 60 were considered significant, and lower-scoring proteins were verified manually or rejected.

Protein correlation profiling

The abundance of each peptide in each of the sucrose gradient fractions was calculated from the area of the extracted ion current peak (see Fig. 1c). To predict elution times in different nanoLC MS/MS experiments, the centroids of the extracted ion currents were determined for sequenced and identified peptides. For several hundred such peptides, the elution times were fitted to the elution time of the identical peptide in the peak centrosomal fraction (fraction 7). From a linear fit in each of the three gradient segments of the LC MS/MS runs, the elution time of the peptides could be predicted within 1% of total gradient time. A program written in the .Net environment (Microsoft) extracted raw mass spectrometric data from PE Sciex data (.wiff) files; it correlated peptides by elution time, charge state and accurate mass and allowed manual inspection for overlapping, co-eluting peaks in these complex peptide mixtures. A consensus fractionation profile was obtained by averaging the normalized peptide abundance profile for all peptides identifying the known centrosomal proteins centrin 2, γ -tubulin and pericentrin. ' χ^2 values' were calculated as the squared deviation of the normalized profile for all peptides divided by the number of data points, for all peptides that had a minimum of four quantification values. The χ^2 value of a protein was determined as the median of its peptide χ^2 values.

Cloning, tagging and expression of fluorescent proteins

cDNAs for candidate centrosomal proteins were obtained from the Kazusa DNA Research Institute, Japan, and the RZPD, Berlin. cDNAs were cloned N- and C-terminally to enhanced green fluorescent protein and tagged with the Myc epitope by using standard cloning procedures. U2OS cells were cultured, transfected and processed for immunofluorescence microscopy as described previously²⁴. Immunofluorescence microscopy was performed with a Zeiss Axioplan II microscope and an Apochromat 63 $\times$ 1.4 n.a. oil-immersion objective. Pictures were taken with a Micromax charge-coupled device camera (Princeton Instruments) and Metaview software (Universal Imaging Corp.), and images were processed with Photoshop (Adobe Systems).

Received 8 August; accepted 30 October 2003; doi:10.1038/nature02166.

- Bornens, M. Centrosome composition and microtubule anchoring mechanisms. *Curr. Opin. Cell Biol.* **14**, 25–34 (2002).
- Doxsey, S. Re-evaluating centrosome function. *Nature Rev. Mol. Cell Biol.* **2**, 688–698 (2001).
- Nigg, E. A. Centrosome aberrations: cause or consequence of cancer progression? *Nature Rev. Cancer* **2**, 815–825 (2002).
- Hinchcliffe, E. H. & Sluder, G. 'It takes two to tango': understanding how centrosome duplication is regulated throughout the cell cycle. *Genes Dev.* **15**, 1167–1181 (2001).
- Rieder, C. L., Faruki, S. & Khodjakov, A. The centrosome in vertebrates: more than a microtubule-organizing center. *Trends Cell Biol.* **11**, 413–419 (2001).
- Aebbersold, R. & Mann, M. Mass spectrometry-based proteomics. *Nature* **422**, 198–207 (2003).
- Wigge, P. A. *et al.* Analysis of the *Saccharomyces* spindle pole by matrix-assisted laser desorption/ionization (MALDI) mass spectrometry. *J. Cell Biol.* **141**, 967–977 (1998).
- Bornens, M. & Moudjou, M. Studying the composition and function of centrosomes in vertebrates. *Methods Cell Biol.* **61**, 13–34 (1999).
- Kellogg, D. R., Moritz, M. & Alberts, B. M. The centrosome and cellular organization. *Annu. Rev. Biochem.* **63**, 639–674 (1994).
- Pietromonaco, S. F., Seluja, G. A., Aitken, A. & Elias, L. Association of 14-3-3 proteins with centrosomes. *Blood Cells Mol. Dis.* **22**, 225–237 (1996).
- Lange, B. M., Bachi, A., Wilm, M. & Gonzalez, C. Hsp90 is a core centrosomal component and is required at different stages of the centrosome cycle in *Drosophila* and vertebrates. *EMBO J.* **19**, 1252–1262 (2000).
- Brown, C. R., Doxsey, S. J., Hong-Brown, L. Q., Martin, R. L. & Welch, W. J. Molecular chaperones and the centrosome. A role for TCP-1 in microtubule nucleation. *J. Biol. Chem.* **271**, 824–832 (1996).
- Gergely, F. *et al.* The TACC domain identifies a family of centrosomal proteins that can interact with microtubules. *Proc. Natl Acad. Sci. USA* **97**, 14352–14357 (2000).
- Gascard, P. *et al.* Characterization of multiple isoforms of protein 4.1R expressed during erythroid terminal differentiation. *Blood* **92**, 4404–4414 (1998).
- Casenghi, M. *et al.* Polo-like kinase 1 regulates Nlp, a centrosome protein involved in microtubule nucleation. *Dev. Cell* **5**, 113–125 (2003).
- Chang, P. & Stearns, T. Delta-tubulin and epsilon-tubulin: two new human centrosomal tubulins reveal new aspects of centrosome structure and function. *Nature Cell Biol.* **2**, 30–35 (2000).
- Klotz, C. *et al.* Parthenogenesis in *Xenopus* eggs requires centrosomal integrity. *J. Cell Biol.* **110**, 405–415 (1990).
- Nakagawa, Y., Yamane, Y., Okanoue, T. & Tsukita, S. Outer dense fiber 2 is a widespread centrosome scaffold component preferentially associated with mother centrioles: its identification from isolated centrosomes. *Mol. Biol. Cell* **12**, 1687–1697 (2001).
- Popovici, C. *et al.* The t(6;8)(q27;p11) translocation in a stem cell myeloproliferative disorder fuses a novel gene, FOP, to fibroblast growth factor receptor 1. *Blood* **93**, 1381–1389 (1999).

- Hearn, T. *et al.* Mutation of ALMS1, a large gene with a tandem repeat encoding 47 amino acids, causes Alström syndrome. *Nature Genet.* **31**, 79–83 (2002).
- Collin, G. B. *et al.* Mutations in ALMS1 cause obesity, type 2 diabetes and neurosensory degeneration in Alström syndrome. *Nature Genet.* **31**, 74–78 (2002).
- Romio, L. *et al.* OFD1, the gene mutated in oral-facial-digital syndrome type 1, is expressed in the metanephros and in human embryonic renal mesenchymal cells. *J. Am. Soc. Nephrol.* **14**, 680–689 (2003).
- Lamond, A. I. & Mann, M. Cell biology and the genome projects—a concerted strategy for characterizing multi-protein complexes using mass spectrometry. *Trends Cell Biol.* **7**, 139–142 (1997).
- Mayor, T., Stierhof, Y. D., Tanaka, K., Fry, A. M. & Nigg, E. A. The centrosomal protein C-Nap1 is required for cell cycle-regulated centrosome cohesion. *J. Cell Biol.* **151**, 837–846 (2000).
- Ohta, T. *et al.* Characterization of Cep135, a novel coiled-coil centrosomal protein involved in microtubule organization in mammalian cells. *J. Cell Biol.* **156**, 87–99 (2002).
- Pfannenschmidt, F. *et al.* Chlamydomonas DIP13 and human NA14: a new class of proteins associated with microtubule structures is involved in cell division. *J. Cell Sci.* **116**, 1449–1462 (2003).
- Kenedy, A. A., Cohen, K. J., Loveys, D. A., Kato, G. J. & Dang, C. V. Identification and characterization of the novel centrosome-associated protein CCCAP. *Gene* **303**, 35–46 (2003).
- Chen, Z., Indjeian, V. B., McManus, M., Wang, L. & Dynlacht, B. D. CP110, a cell cycle-dependent CDK substrate, regulates centrosome duplication in human cells. *Dev. Cell* **3**, 339–350 (2002).
- Yang, J. *et al.* Rootletin, a novel coiled-coil protein, is a structural component of the ciliary rootlet. *J. Cell Biol.* **159**, 431–440 (2002).
- Kilmartin, J. V. Sif1p has conserved centrin-binding sites and an essential function in building yeast spindle pole body duplication. *J. Cell Biol.* **162**, 1211–1221 (2003).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank members of the Center for Experimental Bioinformatics (CEBI) and the Nigg group for critical reading of the manuscript and fruitful discussions and M. Matzner for excellent technical support. We also thank M. J. Pebusque (INSERM U119, Marseille, France) for the kind gift of a cDNA clone of FOP. CEBI is supported by a grant by the Danish Natural Research Foundation. E.A.N. acknowledges support from the Max-Planck Society and the 'Fonds der Chemischen Industrie'.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to E.N. (nigg@biochem.mpg.de) or M.M. (mann@bmb.sdu.dk).

Protein kinase C switches the Raf kinase inhibitor from Raf-1 to GRK-2

Kristina Lorenz, Martin J. Lohse & Ursula Quittner

Institut für Pharmakologie und Toxikologie, Versbacher Strasse 9, D-97078 Würzburg, Germany

Feedback inhibition is a fundamental principle in signal transduction allowing rapid adaptation to different stimuli. In mammalian cells, the major feedback inhibitor for G-protein-coupled receptors (GPCR) is G-protein-coupled receptor kinase 2 (GRK-2), which phosphorylates activated receptors, uncouples them from G proteins and initiates their internalization^{1,2}. The functions of GRK-2 are indispensable and need to be tightly controlled³. Dysregulation promotes disorders such as hypertension⁴ or heart failure⁵. In our search for a control mechanism for this vital kinase, here we show that the Raf kinase inhibitor protein^{6–8} (RKIP) is a physiological inhibitor of GRK-2. After stimulation of GPCR, RKIP dissociates from its known target, Raf-1 (refs 6–8), to associate with GRK-2 and block its activity. This switch is triggered by protein kinase C (PKC)-dependent phosphorylation of the RKIP on serine 153. The data delineate a new principle in signal transduction: by activating PKC, the incoming receptor signal is enhanced both by removing an inhibitor from Raf-1 and by blocking receptor internalization. A physiological role for this mechanism is shown in cardiomyocytes in which the down-regulation of RKIP restrains β -adrenergic signalling and contractile activity.

RKIP is a physiological inhibitor of Raf-1 kinase^{6–8}. Surprisingly,

RKIP dissociates from Raf-1 after activation, apparently liberating RKIP for a novel target^{6–8}. Because previous data suggested crosstalk between Raf-1 and G-protein-coupled receptor kinase 2 (ref. 9), we analysed whether this crosstalk could be mediated by RKIP. We found that purified RKIP bound to purified GRK-2 (data not shown), and we asked whether GRK-2 was a target of RKIP. Purified RKIP inhibited the phosphorylation of rhodopsin with an effector concentration for half-maximum response (EC_{50}) of 460 nM (Fig. 1a), at least one order of magnitude more potent in inhibiting GRK-2 than that of Raf-1 (refs 6, 7). In contrast, RKIP was not effective in inhibiting GRK-5 (data not shown), and RKIP up to 10 μ M also did not inhibit other protein kinases known to phosphorylate GPCRs, such as protein kinase A (PKA) or PKC α and PKC δ (data not shown). RKIP also inhibited by about 70% the agonist-induced phosphorylation of different GPCRs in intact cells such as the α_{1B} -adrenergic receptor¹⁰ or the parathyroid hormone (PTH) receptor¹¹ (Fig. 1b, c). RKIP inhibited phosphorylation of the β_2 -adrenergic receptor, which is mediated by both GRK-2 and PKA¹², by only 50% (Fig. 1d). However, concomitant inhibition of PKA and GRK-2 resulted in an almost complete inhibition of β_2 -adrenergic receptor phosphorylation (Fig. 1d). Taken together, these data show that RKIP specifically inhibits GRK-2-mediated phosphorylation of different receptors *in vitro* and in intact cells.

Several mechanisms are conceivable to explain how RKIP inhibits GRK-2. Because GRK-2 requires free G $\beta\gamma$ subunits for activation, inhibition of GRK-2 by G $\beta\gamma$ -binding proteins seemed feasible¹³. However, RKIP did not interfere with G $\beta\gamma$ -mediated effects (data not shown), excluding a direct RKIP-G $\beta\gamma$ interaction. In contrast, GRK-2 and RKIP formed a stable complex in cells, and stimulation of the β_2 -adrenergic receptor approximately doubled the amount of this complex (Fig. 1e). In further analysing how RKIP interacted with GRK-2, we determined the binding site of RKIP in GRK-2

(Fig. 1f). RKIP interacted with wild-type GRK-2, with GRK-2^{1–485} and with the amino terminus of GRK-2, GRK-2^{1–185} (Fig. 1f, lanes 1, 3 and 4) but RKIP did not bind significantly to the carboxy terminus, GRK-2^{495–689} (Fig. 1f, lane 2), although the GRK-2 proteins were expressed at comparable levels (data not shown). Thus, a binding domain for RKIP is in the N terminus of GRK-2, which is important for GRK-2 activity^{14,15} and for interaction between GRK-2 and its receptor¹⁶.

Phosphorylation affects the interaction of RKIP with Raf-1 (refs 6–8). RKIP was also phosphorylated in intact cells in a PKC-dependent manner upon stimulation of the α_{1B} -adrenergic receptor, the PTH or the β_2 -adrenergic receptor (Supplementary Fig. 1a–c). The PKC phosphorylation site of RKIP was identified as serine 153, because purified PKC α or PKC δ phosphorylated wild-type RKIP but not the mutant RKIP^{S153A} (data not shown), and RKIP^{S153A} was also not significantly phosphorylated in intact cells (Supplementary Fig. 1d, e). Serine 153 was also the PKC-phosphorylation site of native RKIP after stimulation of GPCR as determined by phosphopeptide mapping of trypsin/Arg-C-digested RKIP isolated from cardiomyocytes after stimulation of the β -adrenergic receptor (Supplementary Fig. 1f–i). Together these data demonstrate that different GPCRs stimulate the PKC-dependent phosphorylation of RKIP on serine 153.

Phosphorylation on serine 153 was required for the activity of RKIP as a GRK-2 inhibitor. Whereas wild-type RKIP inhibited agonist-induced phosphorylation of the β_2 -adrenergic receptor by GRK-2, the phosphorylation-deficient RKIP^{S153A} did not (Fig. 2a). In line with this finding, inhibition of β_2 -adrenergic receptor phosphorylation by RKIP in HEK-293 cells led to increased isoprenaline (isoproterenol)-stimulated cyclic AMP (cAMP) concentrations, whereas the phosphorylation-deficient mutant, RKIP^{S153A}, did not significantly alter cAMP generation (Fig. 2b). Similar

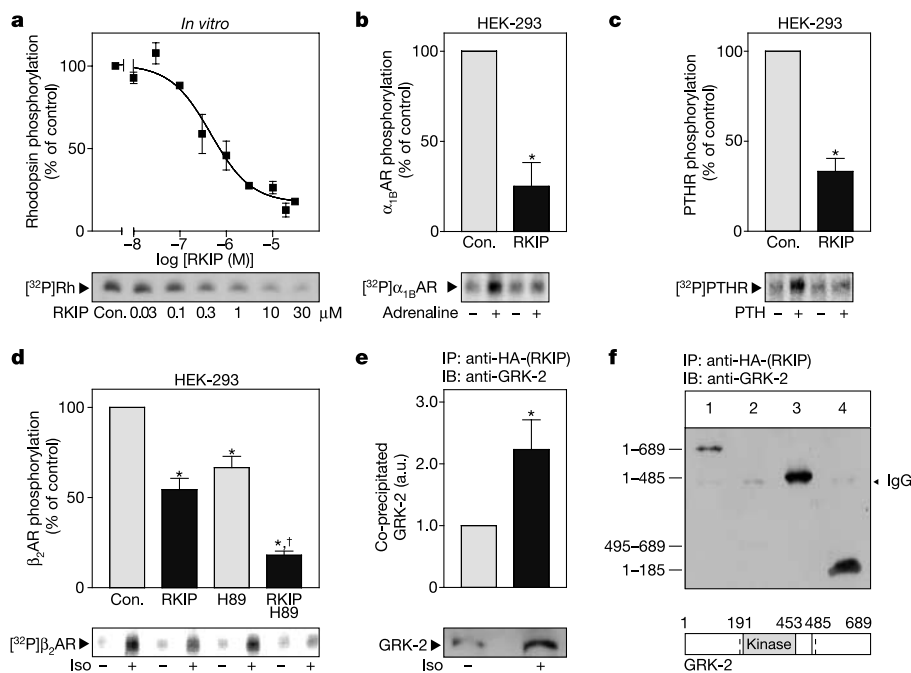


Figure 1 RKIP inhibits GRK-2-mediated GPCR phosphorylation. **a**, Purified RKIP inhibits the phosphorylation of rhodopsin *in vitro* by purified GRK-2. **b–d**, RKIP inhibits the GRK-2-mediated phosphorylation of the α_{1B} -adrenergic receptor (**b**, α_{1B} AR; with or without adrenaline), the PTH receptor (**c**, PTHR; with or without PTH) and the β_2 -adrenergic receptor (**d**, β_2 AR; with or without isoprenaline (Iso)) in ³²P-labelled HEK-293 cells. **e**, Co-immunoprecipitation of GRK-2 with HA-RKIP (IP) from HEK-293 cells expressing HA-RKIP, wild-type GRK-2 and β_2 -adrenergic receptor, and densitometric quantification (arbitrary units (a.u.)) of GRK-2 in immunoblot (IB).

f, Co-immunoprecipitation of GRK-2 proteins with HA-RKIP from HEK-293 cells expressing HA-RKIP together with wild-type GRK-2 (lane 1), GRK-2^{1–485} (lane 2), GRK-2^{1–185} (lane 3) or GRK-2^{495–689} (lane 4). The lower panel shows a topology model of GRK-2. Expression levels of the GRK-2 proteins and of RKIP in the experiment shown in **f**, and the immunoprecipitated RKIP concentrations in **e** and **f**, were similar (data not shown). Data are means $\pm$ s.e.m., $n = 3–6$; asterisk, $P < 0.05$ compared with control (Con.); dagger, $P < 0.05$ compared with inhibition by H89 alone.

findings were obtained with endogenous receptors and GRK-2 of cardiomyocytes (Fig. 2c) and rat osteosarcoma (ROS 17/2.8) cells (Fig. 2d). RKIP also blunted the agonist-stimulated internalization of the β_2 -adrenergic receptor in HEK-293 cells, whereas the phosphorylation-deficient mutant RKIP^{S153A} had no effect (Fig. 2e, f). In conclusion, inhibition of GRK-2 by RKIP enhances signalling and decreases receptor internalization. To act as a GRK-2 inhibitor, RKIP needs to be phosphorylated by PKC on serine 153.

Stimulation of β_2 -adrenergic receptors also activates the Raf-MEK-ERK cascade¹⁷. Coexpression of RKIP further enhanced β_2 -adrenergic receptor-stimulated MAP kinase/ERK kinase (MEK)1/2 activation (data not shown) and extracellular signal-regulated protein kinase (ERK)1/2 activation as determined by immunoblotting (Fig. 2g). Thus, RKIP acts as a GRK-2 inhibitor after stimulation of GPCR, leading to the observed signal enhancement, and upon stimulation of GPCR the inhibition of GRK-2 by RKIP prevails over the inhibition of the Raf-1-MEK-ERK cascade.

To elucidate the mechanism for the apparent predominance of the GRK-2 inhibitory activity of RKIP after stimulation of GPCR, we followed the interaction between GRK-2 and RKIP, and between Raf-1 and RKIP. Stimulation with 12-O-tetradecanoylphorbol-13-acetate (TPA) or PTH increased the amount of GRK-2 co-immunoprecipitated with haemagglutinin (HA)-conjugated RKIP in transfected HEK-293 cells (Fig. 3a, c). In contrast, stimulation with TPA or PTH decreased the amount of Raf-1 that co-immunoprecipitated with HA-RKIP (Fig. 3b, d). The phosphorylation-deficient mutant RKIP^{S153A} remained bound to Raf-1 and did not switch to GRK-2 after stimulation with TPA (data not shown). Thus, phosphorylation of RKIP leads to dissociation of the Raf-1-RKIP complex while stimulating the association of RKIP with GRK-2. The stoichiometry of RKIP phosphorylation after the stimulation of GPCR was compatible with a phosphorylation-regulated switch of RKIP from Raf-1 to GRK-2 (0.15–0.25 mol phosphate per mol RKIP in unstimulated cells increased to 0.4–0.7 mol mol⁻¹ RKIP after stimulation of the α_{1B} -adrenergic or β_2 -adrenergic receptor; data not shown). In addition, Raf-1 kinase activity was correlated with the GPCR-stimulated dissociation of RKIP. Whereas basal Raf-1^{YY340/341DD} activity was suppressed on co-expression of RKIP (Fig. 3e, compare column 3 with column 1), PTH-stimulated Raf-1 activity was not suppressed in RKIP-expressing cells in comparison with PTH-stimulated cells not expressing RKIP (Fig. 3e, columns 2 and 4). Thus, phosphorylation of RKIP after stimulation of GPCR apparently transforms the Raf-1 inhibitor into a GRK-2 inhibitor.

Does the PKC-dependent switch of RKIP from Raf-1 to GRK-2 also occur in native tissue? GRK-2 and Raf-1 are both widely expressed, and high protein expression levels of RKIP were detected in various organs (Fig. 3f). The protein concentrations of RKIP in the different tissues were 5–10-fold that of RKIP in the transiently transfected HEK-293 cells used for the experiments shown above (Fig. 3f). To find out whether the substrate switch of RKIP operated in native tissue, RKIP was isolated from different native systems (brain tissue, ROS 17/2.8 cells and cardiomyocytes), and co-immunoprecipitated GRK-2 or Raf-1 were detected (Fig. 3g). Efficient co-immunoprecipitation of GRK-2 with RKIP occurred only after stimulation of brain tissue, ROS 17/2.8 cells or cardiomyocytes with TPA, PTH or isoprenaline, respectively, whereas high concentrations of the Raf-1-RKIP complex were isolated from unstimulated tissue/cells only (Fig. 3g). Together these results demonstrate that the PKC-dependent switch of RKIP from Raf-1 to GRK-2 also occurs in native tissue and cells.

Are the cellular stoichiometries of RKIP, Raf-1 and GRK-2 sufficient for RKIP to fulfil a physiological role as a bifunctional inhibitor mutually controlling the activities of Raf-1 and GRK-2? In the three different native systems, the stoichiometry of RKIP and Raf-1 was about 1:1, and that of RKIP and GRK-2 was more than 1:0.1. The absolute expression levels of RKIP, Raf-1 and GRK-2 in

brain were 170, 170 and 0.5 pmol per mg cytosolic protein, those in ROS 17/2.8 cells were 160, 170 and 1.6 pmol mg⁻¹, and those in cardiomyocytes were 40, 40 and 0.2 pmol mg⁻¹ (data not shown). Thus, the cellular stoichiometries are sufficient for Raf-1 to bind all RKIP molecules under basal conditions, and for RKIP to bind all GRK-2 molecules on stimulation.

This conclusion was confirmed in cardiomyocytes. Depletion of Raf-1 from unstimulated cells concomitantly decreased RKIP

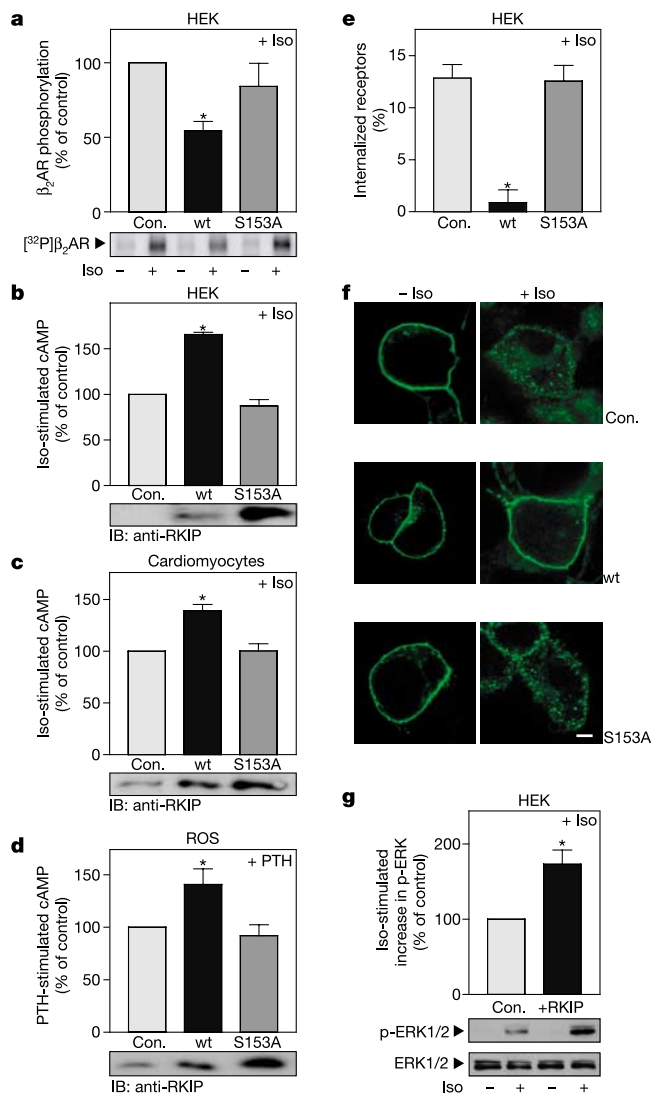


Figure 2 The RKIP-mediated inhibition of GRK-2 correlates with enhanced signalling and decreased internalization. **a**, In contrast to wild-type RKIP (wt), the phosphorylation-deficient RKIP^{S153A} (S153A) did not act as a GRK-2 inhibitor in cells. **b**, Co-expression of wild-type RKIP increased isoprenaline-stimulated cAMP concentrations of HEK-293 cells compared with control cells expressing β_2 -adrenergic receptors only (Con.), whereas the phosphorylation-deficient RKIP^{S153A} (S153A) did not. **c**, **d**, Expression of wild-type RKIP increased receptor-stimulated signalling of native β -adrenergic receptors on cardiomyocytes (**c**) and of PTH receptors on ROS 17/2.8 cells (**d**), but RKIP^{S153A} had no effect. Results in **a–d** are means $\pm$ s.e.m., $n = 3$. **e**, Wild-type RKIP inhibited agonist-induced internalization of β_2 -adrenergic receptors of HEK-293 cells compared with control cells, whereas RKIP^{S153A} (S153A) did not. Results are means $\pm$ s.e.m., $n = 9–13$. **f**, Detection of receptor internalization by fluorescence microscopy with the use of HEK-293 cells expressing GFP-tagged β_2 -adrenergic receptors (control) together with wild-type RKIP or RKIP^{S153A} (S153A). Scale bar, 7 μ m. **g**, RKIP enhances the β_2 -adrenergic receptor-stimulated ERK1/2 activation in HEK-293 cells (means $\pm$ s.e.m., $n = 3$). Asterisk indicates statistical significance ($P < 0.05$ compared with control).

concentrations by about 80%, whereas in TPA-stimulated cells, Raf-1 depletion did not significantly decrease RKIP concentrations (Fig. 3h). Complementary results were obtained for the RKIP-GRK-2 complex of cardiomyocytes. RKIP depletion in TPA-stimulated cells decreased cardiomyocyte GRK-2 by about 80%, whereas under unstimulated conditions RKIP depletion did not significantly affect GRK-2 concentrations (Fig. 3i). Taken together, these findings show that RKIP is mostly bound to Raf-1 under

basal conditions. After stimulation of PKC, phosphorylated RKIP molecules previously bound to Raf-1 were released, and bound and inhibited GRK-2 (Fig. 3j). In this context RKIP seems to be a physiological inhibitor that switches from Raf-1 to GRK-2 after PKC-dependent phosphorylation (Fig. 3j).

Downregulation of endogenous RKIP expression by an RKIP antisense plasmid⁶ or by small interfering RNA (siRNA) duplexes¹⁸ confirmed the physiological role of RKIP as a GRK-2 inhibitor. Gene

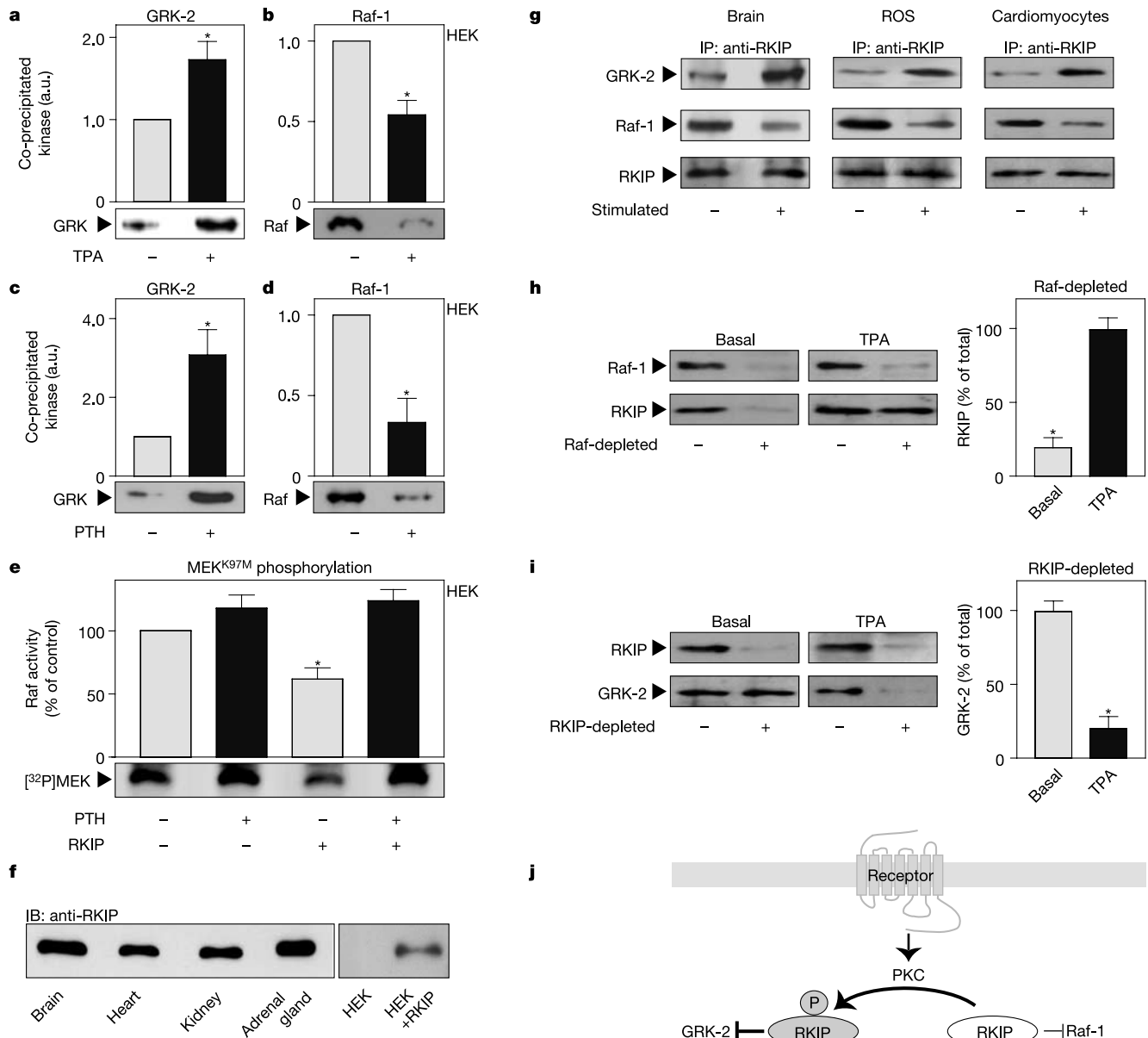


Figure 3 PKC switches RKIP from Raf-1 to GRK-2. **a–d**, HA-RKIP was immunoprecipitated from HEK-293 cells expressing RKIP, GRK-2 and Raf-1 (**a–d**) together with the PTH receptor (**c, d**), and co-immunoprecipitated GRK-2 or Raf-1 were quantified in immunoblot with anti-GRK-2 antibodies (**a, c**), or with anti-Raf-1 antibodies (**b, d**). The amount of immunoprecipitated RKIP was similar in all the experiments (not shown). **e**, RKIP inhibits basal Raf-1^{YY340/341DD} activity of unstimulated cells but not the PTH-stimulated Raf activity. **f**, RKIP detection in mouse brain, heart, kidney and adrenal gland (30 µg protein per lane), and in lysates of control HEK-293 cells or RKIP-expressing HEK-293 cells (50 µg protein per lane). **g**, Immunoprecipitation of RKIP with anti-RKIP antibodies from brain tissue (with or without TPA), from ROS 17/2.8 cells (with or without PTH) or from cardiomyocytes (with or without isoprenaline), and immunoblot determination of the amount of co-immunoprecipitated GRK-2 or Raf-1, or of precipitated

RKIP. Blots are representative of at least five independent experiments, each with similar results. **h**, Immunoblot detection of Raf-1 and RKIP in cytosols of cardiomyocytes before (–) and after (+) depletion of Raf-1 by anti-Raf-1 antibodies from unstimulated cardiomyocytes (left panels, basal) or from TPA-stimulated cells (middle panels, TPA). **i**, Immunoblot detection of RKIP and GRK-2 in cytosols of cardiomyocytes before (–) and after (+) depletion of RKIP from unstimulated cardiomyocytes (left panels) or from TPA-stimulated cells (middle panels). Results are means ± s.e.m., $n = 3–6$; asterisk, $P < 0.05$ compared with control; compared with total in the right panels of **h** and **i**. **j**, Model of RKIP action in cells. RKIP switches from Raf-1 to GRK-2 on PKC phosphorylation. As a bifunctional inhibitor, RKIP suppresses Raf-1 under basal conditions and GRK-2 after receptor stimulation.

silencing of RKIP decreased receptor-stimulated cAMP concentrations in Rat-1, NIH-3T3 and ROS 17/2.8 cells and in cardiomyocytes (Fig. 4a, b; Supplementary Fig. 2a–c). The RKIP antisense and RNAi plasmids selectively downregulated RKIP without affecting GRK-2 expression (Supplementary Fig. 2a–c; Fig. 4a, b) or receptor expression (data not shown). Basal cAMP concentration (Fig. 4b) or forskolin-stimulated cAMP concentrations (data not shown) were not changed on gene silencing. In cardiomyocytes, β -adrenergic receptor-stimulated cardiomyocyte contractile activity decreased concomitantly with cAMP concentration (Fig. 4c). Inhibition of RKIP by RKIP-specific antibodies also reduced β -adrenergic receptor-stimulated beating frequency (compare Fig. 4d with Fig. 4c). In contrast, inhibition of GRK-2 by specific antibodies¹⁵ increased the β -receptor-stimulated response (Fig. 4d), confirming previous results^{1,2}. Interestingly, simultaneous application of RKIP- and GRK-2-specific antibodies increased the β -adrenergic receptor-stimulated contractile activity by an extent similar to the application of the GRK-2-specific antibodies alone (Fig. 4d). Because GRK-2 inhibition neutralized the effect of RKIP inhibition, these results suggest that endogenous RKIP enhances β -adrenergic receptor responsiveness specifically by inhibiting GRK-2.

The GPCR/PKC-regulated switch of a single inhibitor between two different substrates and pathways is without precedent. By switching from Raf-1 to GRK-2, the kinase inhibitor RKIP acts as a signal modifier that (1) enhances receptor signalling by inhibiting GRK-2, and (2) increases the signal:noise ratio of Raf-1-dependent signalling by suppressing basal Raf-1 activity. The substrate switch of RKIP reveals a previously unrecognized signalling network between Raf-1 and GRK-2. This network might help to sustain

the balance between the growth-promoting pathway of Raf-1 and the inhibitory pathway of GRK-2. □

Methods

Plasmids, protein expression and purification

The cDNAs encoding RKIP⁶, HA-RKIP⁶, α_{1B} -adrenergic receptor¹⁰, PTH receptor¹¹, HA-PTH receptor¹¹, Flag- β_2 -adrenergic receptor, β_2 -adrenergic receptor tagged with green fluorescent protein (GFP)¹⁹, GRK-2 (ref. 15), Raf-1 (ref. 9), HA-Raf-1^{YY340/341DD} (ref. 20) and PKC δ (ref. 14) were subcloned into cytomegalovirus promoter-based expression plasmids and used for expression in cells²¹. Gene silencing of RKIP by RNAi was performed by subcloning pairs of oligonucleotides producing a hairpin siRNA and corresponding to positions 537–555 or 152–170 relative to the start codon of RKIP for RNAi-1 or RNAi-2, respectively, into pSilencer 2.0-U6 siRNA expression plasmid¹⁸. Purification of RKIP as a glutathione S-transferase (GST) fusion protein was performed as described¹⁵. All other proteins were expressed in and purified from *Spodoptera frugiperda* (Sf9) cells infected with recombinant baculoviruses encoding PKC α , PKC δ , GRK-2 or His₆-tagged MEK^{K97M} (ref. 20). Mutants and constructs generated by PCR²² were sequenced entirely.

Immunoblot detection of proteins

RKIP, GRK-2 and Raf-1 expression in cells and in tissue were determined in immunoblots^{21,22} with anti-RKIP antibodies raised against recombinant GST-RKIP as antigen, or against a synthetic peptide (sc-5424; Santa Cruz), with anti-GRK-2 antibodies raised against purified human GRK-2 (refs 11, 15) or against a synthetic peptide (sc-562; Santa Cruz), or with anti-Raf-1 antibodies (Transduction Laboratories).

Immunoprecipitation of RKIP

HEK-293 cells were transfected with the indicated cDNAs; 40 h after transfection, cells were stimulated with 1 μ M TPA for 5 min and lysed. Immunoprecipitation of HA-RKIP was performed with anti-HA antibodies adsorbed on Protein A-Sepharose. Co-immunoprecipitated GRK-2 or Raf-1 was detected in an immunoblot with anti-GRK-2 or anti-Raf-1 antibodies. To detect the GRK-2-RKIP or Raf-1-RKIP complex, RKIP was immunoprecipitated with anti-RKIP antibodies from bovine brain lysates, ROS cells or cardiomyocytes pretreated for 5 min with 10 μ M TPA, 1 μ M PTH or 1 μ M isoprenaline as indicated, and co-immunoprecipitated GRK-2 or Raf-1 were quantified^{6,11}.

Phosphorylation assays

The kinase activity of GRK-2 was assessed *in vitro* by phosphorylation of the receptor substrate rhodopsin¹⁵. Phosphorylation of GPCRs and of RKIP was determined in intact [³²P]orthophosphate-labelled cells as described¹¹. Inhibition of Raf-1 kinase activity by RKIP was assessed after immunoprecipitation of HA-Raf-1^{YY340/341DD} by an *in vitro* phosphorylation assay with MEK^{K97M} as substrate^{9,20}. The phosphorylated proteins were separated by SDS-polyacrylamide gel electrophoresis and analysed by autoradiography or PhosphorImager analysis.

Determination of the stoichiometry of RKIP phosphorylation

The stoichiometry of RKIP phosphorylation was determined in RKIP-expressing [³²P]orthophosphate-labelled HEK-293 cells as described²³. The amount of [³²P]orthophosphate incorporated into the cellular ATP pool was determined by autoradiography after separation of cellular nucleotides by thin-layer chromatography, and total cellular ATP concentrations were determined by luciferase assay (Promega). Immunoprecipitated RKIP concentrations were quantified in the immunoblot, and the amount of [³²P]ATP incorporated into RKIP was determined by PhosphorImager analysis with [γ -³²P]ATP as standard.

Two-dimensional phosphopeptide mapping of RKIP from cardiomyocytes

Details are provided in Supplementary Information.

Determination of cellular cAMP concentrations

Cellular cAMP concentrations were determined to assess the effect of RKIP on β_2 -adrenergic or PTH receptor signalling. HEK-293 cells were transfected with the β_2 -adrenergic receptor and GRK-2. Transfection of additional plasmids was performed as indicated. Cells were stimulated at 37 °C with 1 nM–1 μ M isoprenaline (HEK-293 cells and cardiomyocytes) or with 100 nM PTH^{1–34} (ROS 17/2.8 cells) for 5 min in the presence of 100 μ M 3-isobutyl-1-methylxanthine, and cellular cAMP was determined by radioimmunoassay²⁴.

Isolation of rat cardiomyocytes and contractile activity

Neonatal rat cardiomyocytes were isolated as described²⁵. Cardiomyocytes were transfected by electroporation²⁶ (transfection efficiency more than 60%). Electroporation was also used to load cardiomyocytes with affinity-purified (100–200 nM) RKIP-specific antibodies (sc-5424) or GRK-2-specific antibodies crossreactive with the N terminus of GRK-2 (ref. 15) to inhibit RKIP⁶ or GRK-2 activity¹⁵, respectively.

To monitor the contractile activity of cardiomyocytes, live cells were examined on a Zeiss Axiovert 135 microscope. Cells held thermostatically at 31 °C were observed in MEM²⁵, and the number of beats per minute was determined before and 2 min after stimulation with 5 nM isoprenaline by an observer who was blind to the experimental protocol studied.

β_2 -adrenergic receptor internalization and fluorescence microscopy

Internalization of β_2 -adrenergic receptors expressed in HEK-293 cells was determined by binding of [³H]CGP-12177 as described²⁷.

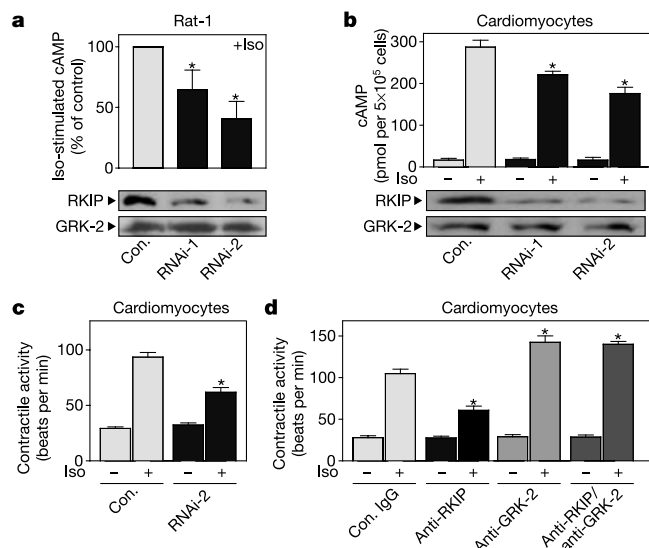


Figure 4 Endogenously expressed RKIP enhances receptor-stimulated signalling. **a**, Gene silencing of RKIP in Rat-1 cells by two different RNAi plasmids, RNAi-1 and RNAi-2, decreased isoprenaline-stimulated cAMP concentrations compared with cells transfected with an unrelated control RNAi plasmid (Con., set to 100%). RKIP and GRK-2 protein concentrations were controlled in the immunoblot (lower panels). **b**, Gene silencing of RKIP in cardiomyocytes by transfection of plasmids RNAi-1 and RNAi-2. The upper panel shows basal and isoprenaline-stimulated cAMP concentrations, and the lower panels show immunoblots of RKIP and GRK-2 protein concentrations. **c**, Gene silencing of RKIP by RNAi-2 decreased isoprenaline-stimulated contractile activity. **d**, Contractile activity of cardiomyocytes treated with control IgG (Con. IgG), with RKIP-specific antibodies, with GRK-2-specific antibodies (anti-GRK-2), or with RKIP- and GRK-2-specific antibodies (anti-RKIP/anti-GRK-2). Results are means $\pm$ s.e.m., $n = 3$ –6; asterisk, $P < 0.05$ compared with control.

For internalization of GFP-tagged β_2 -adrenergic receptors, transfected HEK-293 cells were grown on glass coverslips and, 24 h after transfection, were stimulated with isoprenaline. Receptor internalization was followed by fluorescence microscopy with a Leica (TCS) confocal laser microscope after fixing the cells with 4% (w/v) paraformaldehyde¹⁹.

Statistical analysis

Statistical significance was evaluated by unpaired or paired *t*-tests, and by analysis of variance as applicable; *P* < 0.05 was considered statistically significant.

Received 27 August; accepted 10 October 2003; doi:10.1038/nature02158.

1. Krupnick, J. G. & Benovic, J. L. The role of receptor kinases and arrestins in G protein-coupled receptor regulation. *Annu. Rev. Pharmacol. Toxicol.* **38**, 289–319 (1998).
2. Lefkowitz, R. J., Pitcher, J., Krueger, K. & Daaka, Y. Mechanisms of β -adrenergic receptor desensitization and resensitization. *Adv. Pharmacol.* **42**, 416–420 (1998).
3. Jaber, M. *et al.* Essential role of β -adrenergic receptor kinase 1 in cardiac development and function. *Proc. Natl Acad. Sci. USA* **93**, 12974–12979 (1996).
4. Gros, R., Benovic, J. L., Tan, C. M. & Feldman, R. D. G-protein-coupled receptor kinase activity is increased in hypertension. *J. Clin. Invest.* **99**, 2087–2093 (1997).
5. Ungerer, M., Böhm, M., Elce, J. S., Erdmann, E. & Lohse, M. J. Altered expression of β -adrenergic receptor kinase and β_2 -adrenergic receptors in the failing human heart. *Circulation* **87**, 454–463 (1993).
6. Yeung, K. *et al.* Suppression of Raf-1 kinase activity and MAP kinase signalling by RKIP. *Nature* **401**, 173–177 (1999).
7. Yeung, K. *et al.* Mechanism of suppression of the Raf/MEK/extracellular signal-regulated kinase pathway by the raf kinase inhibitor protein. *Mol. Cell. Biol.* **20**, 3079–3085 (2000).
8. Corbit, K. C. *et al.* Activation of Raf-1 signaling by protein kinase C through a mechanism involving Raf kinase inhibitory protein. *J. Biol. Chem.* **278**, 13061–13068 (2003).
9. Slupsky, J. R. *et al.* Binding of G $\beta\gamma$ subunits to cRaf downregulates G-protein-coupled receptor signalling. *Curr. Biol.* **9**, 971–974 (1999).
10. Diviani, D. *et al.* Effect of different G protein-coupled receptor kinases on phosphorylation and desensitization of the α_{1B} -adrenergic receptor. *J. Biol. Chem.* **271**, 5059–5058 (1996).
11. Dicker, F., Quitterer, U., Winstel, R., Honold, K. & Lohse, M. J. Phosphorylation-independent inhibition of parathyroid hormone receptor signaling by G protein-coupled receptor kinases. *Proc. Natl Acad. Sci. USA* **96**, 5476–5481 (1999).
12. Lohse, M. J., Benovic, J. L., Caron, M. G. & Lefkowitz, R. J. Multiple pathways of rapid β_2 -adrenergic receptor desensitization. Delineation with specific inhibitors. *J. Biol. Chem.* **265**, 3202–3211 (1990).
13. Pitcher, J. A. *et al.* Role of $\beta\gamma$ subunits of G proteins in targeting the β -adrenergic receptor kinase to membrane-bound receptors. *Science* **257**, 1264–1267 (1992).
14. Krasel, C. *et al.* Phosphorylation of GRK2 by protein kinase C abolishes its inhibition by calmodulin. *J. Biol. Chem.* **276**, 1911–1915 (2001).
15. Eichmann, T. *et al.* The amino-terminal domain of G-protein-coupled receptor kinase 2 is a regulatory G $\beta\gamma$ binding site. *J. Biol. Chem.* **278**, 8052–8057 (2003).
16. Yu, Q. M. *et al.* The amino terminus with a conserved glutamic acid of G protein-coupled receptor kinases is indispensable for their ability to phosphorylate photoactivated rhodopsin. *J. Neurochem.* **73**, 1222–1227 (1999).
17. Maudsley, S. *et al.* The β_2 -adrenergic receptor mediates extracellular signal-regulated kinase activation via assembly of a multi-receptor complex with the epidermal growth factor receptor. *J. Biol. Chem.* **275**, 9572–9580 (2000).
18. Hasuwa, H., Kaseda, K., Einarsdottir, T. & Okabe, M. Small interfering RNA and gene silencing in transgenic mice and rats. *FEBS Lett.* **532**, 227–230 (2002).
19. Kallal, L., Gagnon, A. W., Penn, R. B. & Benovic, J. L. Visualization of agonist-induced sequestration and down-regulation of a green fluorescent protein-tagged β_2 -adrenergic receptor. *J. Biol. Chem.* **273**, 322–328 (1998).
20. Mason, C. S. *et al.* Serine and tyrosine phosphorylations cooperate in Raf-1, but not B-Raf activation. *EMBO J.* **18**, 2137–2148 (1999).
21. AbdAlla, S., Lother, H. & Quitterer, U. AT₁-receptor heterodimers show enhanced G-protein activation and altered receptor sequestration. *Nature* **407**, 94–98 (2000).
22. AbdAlla, S., Lother, H., el Massiery, A. & Quitterer, U. Increased AT₁ receptor heterodimers in preeclampsia mediate enhanced angiotensin II responsiveness. *Nature Med.* **7**, 1003–1009 (2001).
23. Darman, R. B. & Forbush, B. A regulatory locus of phosphorylation in the N terminus of the Na-K-Cl cotransporter, NKCC1. *J. Biol. Chem.* **277**, 37542–37550 (2002).
24. AbdAlla, S., Lother, H., Abdel-tawab, A. M. & Quitterer, U. The angiotensin II AT₂ receptor is an AT₁ receptor antagonist. *J. Biol. Chem.* **276**, 39721–39726 (2001).
25. Ritter, O. *et al.* AT₂ receptor activation regulates myocardial eNOS expression via the calcineurin-NFAT pathway. *FASEB J.* **17**, 283–285 (2003).
26. Wu, J. C., Tsai, R. Y. & Chung, T. H. Role of catenins in the development of gap junctions in rat cardiomyocytes. *J. Cell. Biochem.* **88**, 823–835 (2003).
27. Dunigan, C. D., Curran, P. K. & Fishman, P. H. Detection of β -adrenergic receptors by radioligand binding. *Methods Mol. Biol.* **126**, 329–343 (2000).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank W. Kolch for RKIP cDNA plasmids, M. Hoffmann for technical assistance, C. Dees for purification of GRK-2, M. Philipp for help with mouse tissue preparation, H. Mischak for PKC α and PKC δ , N. Burkard and S. Oberdorf-Maass for assistance in preparing cardiomyocytes, and S. Freund for determination of cardiomyocyte beating frequency. This work was supported by the Deutsche Forschungsgemeinschaft.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence should be addressed to U.Q. (toph029@rzbox.uni-wuerzburg.de).

Release of eIF6 (p27^{BBP}) from the 60S subunit allows 80S ribosome assembly

Marcello Ceci^{1,2}, Cristina Gaviraghi^{1,2}, Chiara Gorriini^{1,2}, Leonardo A. Sala^{1,2}, Nina Offenhäuser³, Pier Carlo Marchisio^{1,2} & Stefano Biffo^{1,4}

¹Molecular Histology Unit, DIBIT-HSR, ²University 'Vita-Salute' San Raffaele School of Medicine, 20132 Milano, Italy

³Firc Institute for Molecular Oncology, 20100 Milano, Italy

⁴Department of Science and Advanced Technology, University of Eastern Piedmont 'Amedeo Avogadro', 15100 Alessandria, Italy

The assembly of 80S ribosomes requires joining of the 40S and 60S subunits, which is triggered by the formation of an initiation complex on the 40S subunit¹. This event is rate-limiting for translation², and depends on external stimuli³ and the status of the cell⁴. Here we show that 60S subunits are activated by release of eIF6 (also termed p27^{BBP})^{5,6}. In the cytoplasm, eIF6 is bound to free 60S but not to 80S. Furthermore, eIF6 interacts in the cytoplasm with RACK1⁷, a receptor for activated protein kinase C (PKC). RACK1 is a major component of translating ribosomes, which harbour significant amounts of PKC. Loading 60S subunits with eIF6 caused a dose-dependent translational block and impairment of 80S formation, which were reversed by expression of RACK1 and stimulation of PKC *in vivo* and *in vitro*. PKC stimulation led to eIF6 phosphorylation, and mutation of a serine residue in the carboxy terminus of eIF6 impaired RACK1/PKC-mediated translational rescue. We propose that eIF6 release regulates subunit joining, and that RACK1 provides a physical and functional link between PKC signalling and ribosome activation.

eIF6 is partly nucleolar *in vivo*, and its depletion impairs the biogenesis of the 60S ribosomal subunit^{8–10}. However, eIF6 is puzzling in that it keeps the 40S and 60S subunits dissociated *in vitro*. This finding led to the proposal that eIF6 is an initiation factor¹¹. Its role in translation is not known, and so we addressed this issue in mammalian cells. Profiles of soluble cell extracts show that endogenous eIF6 is either free or bound to free 60S subunits but not to 80S and polysomes (several ribosomes) (Supplementary Fig. 1a). In addition, endogenous eIF6 is found mainly in the cytoplasm⁵ (Supplementary Fig. 1b), where it partitions (when analysed by gel filtration) into two equally represented forms, one that is monomeric (relative molecular mass, $M_r = 2.62 \times 10^3$) and one with a higher molecular mass ($V_o \geq 1.3 \times 10^6 M_r$; ribosome $M_r = 2.82 \times 10^6$), which suggests that eIF6 associates with cytoplasmic ribosomes (Supplementary Fig. 1c).

To identify factors that directly interact with eIF6, we performed yeast two-hybrid screening, which led to the isolation of RACK1, a receptor for PKC^{12–14} and a scaffold protein¹⁵. In subsequent studies in A431 cells, endogenous eIF6 co-immunoprecipitated with endogenous RACK1 (Fig. 1a), and affinity purification of extracts with maltose-binding protein (MBP)–RACK1 resulted in enrichment of endogenous eIF6 (Fig. 1b). Far-western experiments showed that recombinant eIF6 directly bound immobilized RACK1 (Fig. 1c). Thus, eIF6 can directly bind RACK1 *in vivo*.

Immunofluorescence studies showed that endogenous RACK1 is present in the cytoplasm but not in the nucleus (Fig. 1d and Supplementary Fig. 2); eIF6 was detected in both compartments^{5,8} (Fig. 4 and data not shown). Overall, the data suggest that the interaction between RACK1 and eIF6 takes place in the cytoplasm. The presence of eIF6 on ribosomes and its ability to bind RACK1 led us to investigate RACK1–ribosome association. Polysomal analysis

showed that endogenous RACK1 was present on ribosomes and polysomes, as well as in the cytoskeletal fraction (Fig. 1e). In the same gradient, eIF6 was found in the cytoskeletal fraction (on intermediate filaments, as in ref. 5), in the free fraction and on 60S ribosomes (Fig. 1e). In purified native ribosomal subunits from HeLa cells, RACK1 seemed to be associated mainly with 40S subunits (Supplementary Fig. 2b–c) and was also enriched in rabbit reticulocyte lysates (Fig. 2d). Thus, RACK1 and eIF6 are present on ribosomes, but their slightly different locations suggest that they have different roles in translation. Furthermore, data from polysomal profiles and immunoprecipitation experiments indicate that binding of RACK1 to eIF6 occurs either transiently on cytoplasmic ribosomes or in the cytoskeletal fraction. Because the buffers used in immunoprecipitation (Fig. 1a) do not solubilize cytoskeleton-

bound eIF6, we favour the interaction with ribosomes.

To understand the physiological significance of the RACK1–eIF6 interaction, the two factors were added to translation-competent reticulocyte extracts. Reticulocyte extracts contained low, but detectable, levels of endogenous eIF6, suggesting a substoichiometric ratio with 60S (Fig. 2). Exogenous messenger RNA for luciferase was added, and luciferase activity was measured. Addition of eIF6 led to a dose-dependent block in translation (Fig. 2a), which occurred at approximately 1:1 stoichiometry with 60S. This block in translation was caused by impairment of 80S formation (Fig. 2b). By contrast, addition of RACK1 to reticulocyte lysates, which are rich in endogenous RACK1 (Supplementary Fig. 2d), neither affected the basal rate of translation (without eIF6) nor reversed eIF6 inhibition.

The effect of these factors was then examined *in vivo*. eIF6 overexpression reduced translation in two reporter assays. First, we transfected COS cells with eIF6 under the control of the SV40 promoter, and, after 36 h, we measured translation using a pulse of radioactive methionine. COS cells transfected with eIF6 had an approximately 60% reduction in translation (Fig. 2c). Second, we obtained the same results by transfecting HeLa Tet-off cells using inducible luciferase as a reporter gene (Fig. 2d). These results are consistent with the effects of eIF6 on translation *in vitro*, with its localization on the 60S subunit *in vivo*, and with the anti-association activity observed *in vitro*^{6,16}. The *in vivo* effect of RACK1 was then tested. Transfection of RACK1 stimulated both translation of the luciferase reporter (Fig. 2d) and the basal rate of translation, as measured by methionine (Fig. 3c and data not shown). We suggest that, because eIF6 binds to mature 60S, because it is absent from

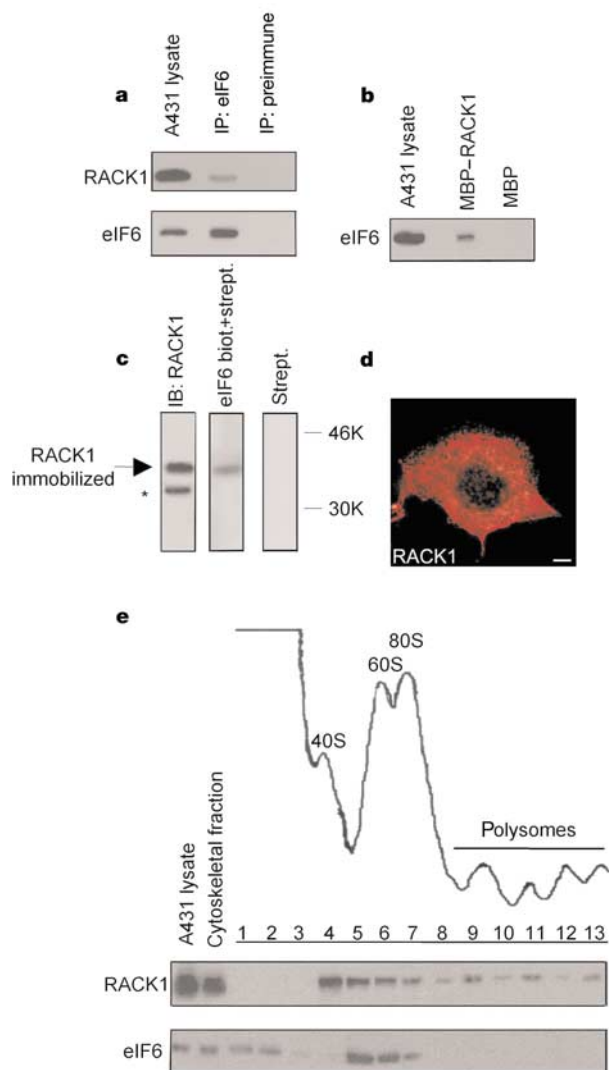


Figure 1 Endogenous RACK1 is found on ribosomes and binds to eIF6 in the cytoplasm. **a**, Interaction between endogenous proteins: A431 lysates immunoprecipitated with anti-eIF6 or preimmune sera were immunoblotted with anti-RACK1 (top panel) and anti-eIF6 (bottom panel). **b**, MBP–RACK1 pulls down endogenous eIF6. MBP–RACK1 and MBP affinity purification followed by immunoblotting for eIF6. **c**, Direct interaction between RACK1 and eIF6 by far-western blotting. Immobilized RACK1 revealed with anti-RACK1 (positive control) or biotinylated eIF6 or streptavidin (negative control). *, degradation product. **d**, **e**, Endogenous RACK1 is present in the cytoplasm by immunofluorescence (**d**) and on ribosomes when detected by immunoblotting of fractions collected by sucrose density gradient (**e**). 1, 2, soluble; 4, 40S peak; 9–13, polysomes. Lower panel, eIF6 (transfected + endogenous). Bar, 5 μ m.

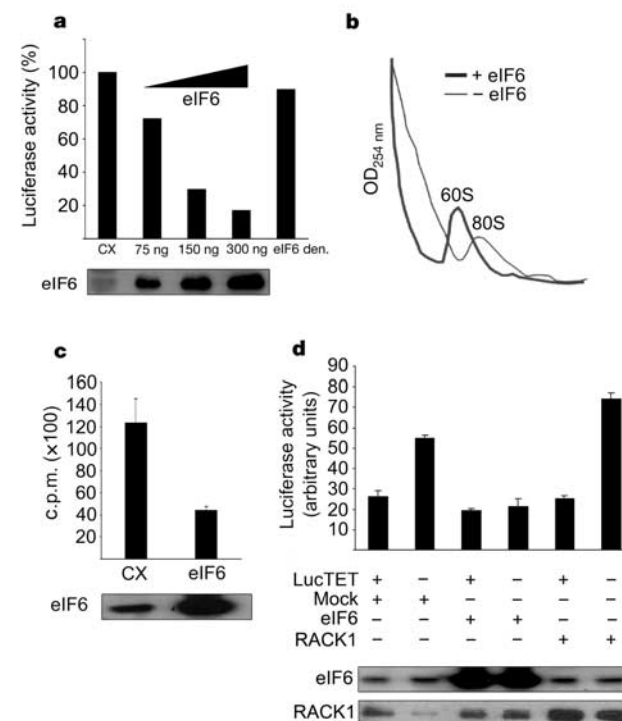


Figure 2 eIF6 blocks translation and RACK1 stimulates translation *in vivo*. **a**, Translation of luciferase mRNA incubated with reticulocyte extracts after adding increasing amounts of recombinant eIF6, no factor (CX) or heat-inactivated eIF6. s.d. is less than 5%. Lower panel, western blotting for eIF6. **b**, Sucrose density gradient of extracts (as in **a**) incubated with or without eIF6. The 40S peak is covered by the soluble fraction. **c**, Translation rate measured by a pulse of radioactive methionine in COS cells transfected with either eIF6 or mock (CX). Lower panel, western blotting for eIF6, 10 mg total protein. **d**, Luciferase activity in cells co-transfected with inducible luciferase and the indicated factors. Lower panels, western blotting for eIF6 and RACK1, as in **c**.

translating polysomes, and because it represses translation, a mechanism for its release must exist. RACK1 might be involved in this release because it binds eIF6 and is found on polysomes. However, as binding of RACK1 to eIF6 cannot modulate either the activity of eIF6 or the rate of translation *in vitro*—although *in vivo* it does have a stimulatory effect on translation—we assume that additional factors are involved.

Next, we tested whether PKC was located on translating ribosomes. Selected PKC isoforms were found to associate with ribosomal fractions (Fig. 3a; L.A.S., P.C.M. and S.B., unpublished data). The presence of PKC on translating ribosomes was further assessed by precipitation of poly(A) mRNA with oligodT, and we assumed that translating ribosomes would be pulled down, but free 60S would not. OligodT precipitation led to enrichment in RACK1 and in PKC β II. Under these conditions, eIF6 was not detected, confirming its exclusion from polysomes (Fig. 3b). To test whether PKC is involved in the interaction of eIF6 with RACK1, the effect of PKC stimulation on the translational inhibition exerted by eIF6 was studied (Fig. 3c). We used serum-grown cells to prevent growth-factor deprivation, which could lead to translational repression. Under these conditions, PKC stimulation did not increase translation in mock-transfected cells (Supplementary Fig. 3a). The results indicated: first, that eIF6 alone repressed translation, and that coexpression of RACK1 led to a minor recovery; and second, that stimulation of PKC activity by 12-*O*-tetradecanoylphorbol-13-acetate (TPA) resulted in recovery of the eIF6 translational block. In

addition, treatment with PKC inhibitors in cells overexpressing eIF6 abolished the increase in translation derived from both TPA and epidermal growth factor, suggesting a specific involvement of the PKC pathway (Supplementary Fig. 3b–d). In agreement with metabolic-labelling data, cells transfected with eIF6 had a large amount of free 60S, and after stimulation of PKC, they showed 80S polysome formation (Fig. 3d), which was associated with release of eIF6 (Supplementary Fig. 4).

We proposed that a direct effect of RACK1 and PKC on eIF6 activity could be associated with increased phosphorylation of endogenous eIF6, and that phosphorylated eIF6 could not inhibit translation *in vitro*. Cells were metabolically labelled with 32 P, then immunopurified for eIF6 and analysed for phosphorylated eIF6 in the cytosol. PKC stimulation increased the amount of phosphorylated eIF6 (Fig. 3e). These data are consistent with a model in which the RACK1–PKC complex rescues translation either by directly modifying eIF6 or by acting on the ribosomal machinery as a whole, or both. The part played by eIF6 phosphorylation in repressing translation was assessed *in vitro*. The ability of PKC to phosphorylate eIF6 was examined first. Recombinant eIF6 was phosphorylated by the RACK1–PKC β II complex (Fig. 3f), but with low stoichiometry (a maximum of 15–20% input of eIF6). The effect of phosphorylated eIF6 on translational activity *in vitro* was then examined as in Fig. 2a. Phosphorylated eIF6 had less anti-translational activity than the non-phosphorylated protein (Fig. 3g). In this specific assay, direct addition of either activated PKC or RACK1 to lysates had no effect (data not shown).

Release of eIF6 by phosphorylation is predicted to occur in the cytoplasm, where RACK1 and PKC β II reside. TPA-stimulated cells were labelled with 32 P, nuclear and cytoplasmic extracts were fractionated, and endogenous eIF6 was immunopurified. Phosphorylated eIF6 was localized mainly in the cytoplasm where the overall amount of eIF6 was higher (Fig. 4a). We tried to establish which eIF6 residues are phosphorylated *in vivo*. Double-tagged eIF6 was transfected into COS cells, which were lysed after PKC

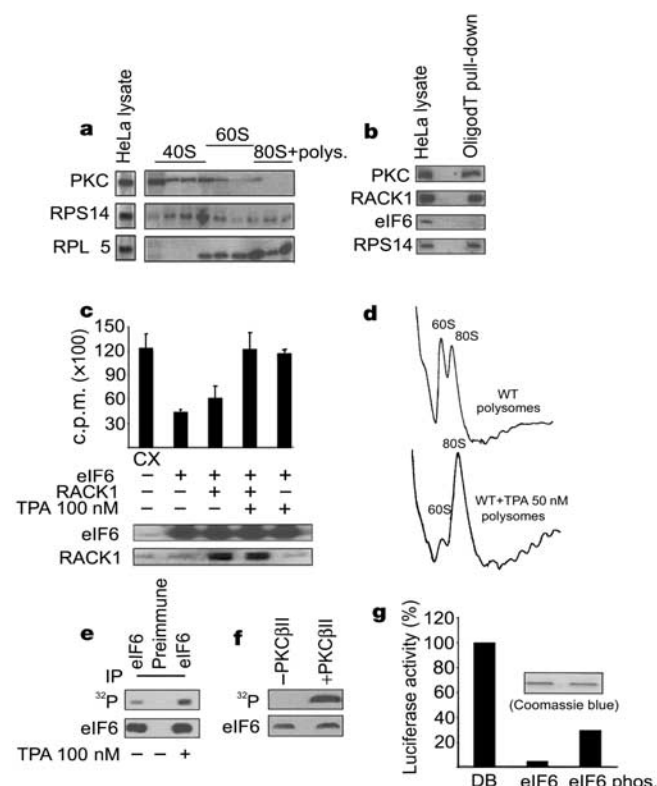


Figure 3 The RACK1–PKC complex acts on eIF6. **a**, **b**, PKC is present on translating ribosomes. **a**, Ribosomal fractions immunoblotted for PKC and polysomal markers. The soluble fraction is omitted. **b**, OligodT pull-down for poly(A) RNA, followed by immunoblotting for the indicated proteins, PKC β II isoform. **c**, PKC stimulation rescues eIF6 inhibition. Translation rate measured as in Fig. 2c of cells transfected with various factors and stimulated with TPA. Lower panels, western blotting. **d**, Ribosomal profile of cells transfected with eIF6 with (below) or without (top) PKC stimulation showing 80S polysome formation. **e**, Phosphorylation of eIF6 *in vivo* after stimulation of PKC (as in Fig. 4d). **f**, eIF6 phosphorylated *in vitro* by RACK1–PKC β II. **g**, Phosphorylated eIF6 does not inhibit translation *in vitro* (see Methods and main text). Data are shown as a percentage of the controls; $n = 7$. DB, dialysis buffer. Inset shows the levels of eIF6.

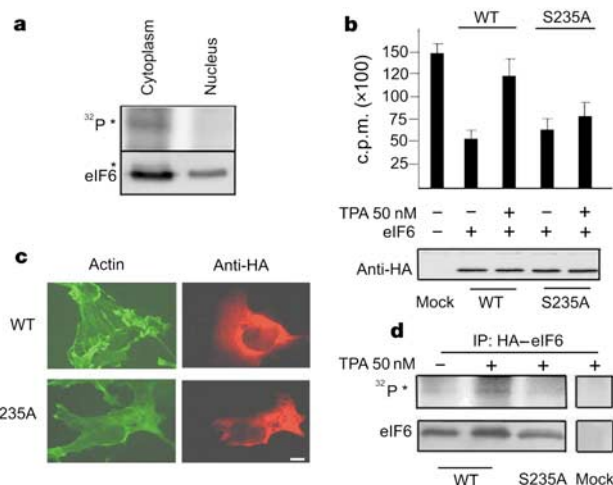


Figure 4 Reduction of anti-association activity by phosphorylation *in vivo*. **a**, Cytoplasmic eIF6 is phosphorylated. Extracts of 32 P-labelled cells, immunoprecipitated with eIF6 antibodies, subjected to electrophoresis, transferred and autoradiographed. Upper panel, autoradiography; lower panel, immunoblotting. *, phosphorylated form. **b**, Translation of the S235A eIF6 mutant is not affected by PKC stimulation. The translation rate was measured as in Fig. 3c. Lower panel, immunoblotting for transfected eIF6. **c**, Wild-type and mutant eIF6 are both in the cytoplasm. Immunofluorescence for indicated proteins; actin staining revealed by fluorescein isothiocyanate (FITC)-labelled phalloidin. Bar, 4 μ m. **d**, Phosphorylation of transfected factors. Transfected eIF6 was sequentially purified with anti-HA and anti-eIF6 antibodies, and purified proteins were analysed by electrophoresis, transfer and autoradiography. Upper panel, 32 P autoradiography; lower panel, immunoblotting.

stimulation. eIF6 was immunopurified to homogeneity and subjected to mass spectrometric analysis (Supplementary Methods). No clear phosphopeptides were found, suggesting low stoichiometry and/or poor recovery of specific peptides. We reasoned that a functional analysis of eIF6 would give a clue to the phosphorylation site *in vivo*. Deletion constructs were screened for their inability to stimulate translation in a TPA-regulated fashion. A construct with a 23-amino-acid deletion at the C terminus of eIF6 could not rescue translation upon PKC stimulation (data not shown). In this region, only one serine (Ser 235) embedded in a low-stringency consensus sequence (IATSMRD) for PKC phosphorylation was observed. This serine was mutated to alanine, and the resulting S235A mutant eIF6 was transfected into COS cells. The increase in translation upon PKC stimulation was tested using wild-type eIF6 and the S235A mutant. Data showed that the S235A mutant repressed translation and could not fully rescue it upon PKC stimulation (Fig. 4b). In agreement with this, overexpression of the S235A mutant caused flat polysomes and 60S accumulation, which were not significantly rescued by PKC stimulation (Supplementary Fig. 5). The S235A mutant was correctly localized as the wild-type eIF6 in both the nucleus and the cytoplasm (Fig. 4c). The extent of TPA-stimulated incorporation of radioactive phosphate was then addressed, comparing wild-type eIF6 and S235A. Increased incorporation of ^{32}P was observed in wild-type eIF6 but not in the S235A mutant (Fig. 4d and Supplementary Fig. 6). Background levels of ^{32}P incorporation were seen in both the wild type and the S235A mutant and may indicate the presence of additional phospho-sites (Supplementary Fig. 6).

To evaluate whether translational inhibition could be linked to 80S inhibition, wild-type eIF6 and S235A eIF6 were studied in an *in vitro* system similar to the one where eIF6 activity was first

shown¹¹. Salt-washed purified 40S and 60S ribosomal subunits form 80S following a sharp increase in Mg^{2+} from 1 to 5 mM. This association is inhibited by eIF6 (Fig. 5a–c and ref. 11), which binds to the 60S subunit and prevents 80S formation. Addition of activated PKC β II or of RACK1–PKC β II reduced the inhibition of 80S formation due to wild-type eIF6 inducing its release from the 60S subunit (Fig. 5c–e). S235A eIF6 bound the 60S subunit, inhibiting 80S formation in the presence and absence of RACK1–PKC β II (Fig. 5f, g). Salt-washed purified subunits contained low levels of endogenous RACK1 and detectable eIF6, and were also able to bind exogenously added factors (Fig. 5a–g). Within the limits of this technique, exogenous RACK1 could be seen both bound to the 40S subunit and in the soluble fraction. In addition, RACK1 was retained in the 80S, where eIF6 was absent.

Our results show that release of eIF6 from 60S subunits may operate, in mammalian cells, through a RACK1–PKC β II pathway. Do other pathways exist? In yeast cells, release of eIF6 from 60S is partially mediated through the efl1 protein¹⁷. No functional data are available about mammalian efl1p. Concerning RACK1, the yeast homologue, Asc1p, is in the eIF6 complex (M.C. *et al.*, unpublished data), on ribosomes (<http://yeast.cellzome.com/>), and ASC1 mutants have impaired 80S formation¹⁸. Although this suggests evolutionary conservation of RACK1-mediated 60S activation, PKC isoforms and pathways in yeasts and mammals diverge. This might be because in yeasts, RACK1 has a function that is independent of PKC, or because the RACK1 and efl1p pathways independently coexist, modulating the binding affinity of eIF6 for 60S. In addition, mammalian cells may have a RACK1–PKC ribosomal pathway. In mammalian cells, where integration of external stimuli with the translational machinery is required^{3,19,20}, the localization of the RACK1–PKC β II complex on ribosomes can bring activated PKC

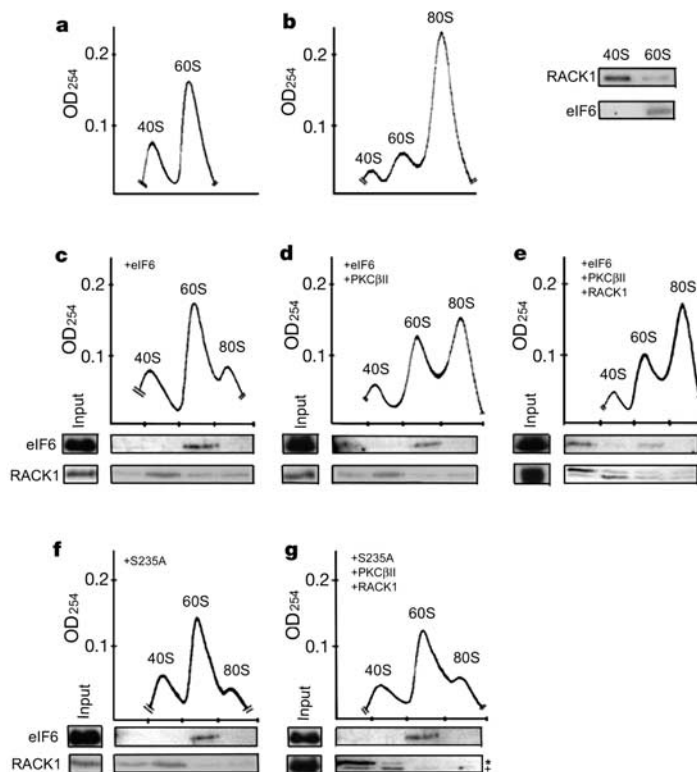


Figure 5 eIF6 anti-association activity *in vitro* is modulated by RACK1–PKC β II. 0.7 pmol 80S was used per experiment. Four fractions were assayed: soluble, 40S, 60S and 80S. **a, b**, Salt-washed ribosomal subunits remain dissociated in 1 mM Mg^{2+} (**a**) but associate after raising the concentration to 5 mM Mg^{2+} , in the absence of eIF6 (**b**, left panel). The right panel of **b** shows endogenous proteins in salt-washed ribosomes.

c, Association of ribosomal subunits in 5 mM Mg^{2+} is inhibited by eIF6 (top), which associates with the free 60S fraction (lower panel). **d, e**, Addition of activated PKC β II (**d**) or RACK1–PKC β II (**e**) results in release of eIF6 from the 60S subunit (lower panel, soluble fraction) and 80S joining (top). **f, g**, S235A eIF6 inhibits ribosomal joining in the presence of RACK1–PKC β II. The blots also show RACK1 (see main text for details).

close to relevant substrates. Apart from eIF6, whose release from 60S may allow the final maturation of 60S in a pioneer round of translation, other substrates may exist. In the past, eIF4G was reported to provide a scaffold for Mnk kinase on eIF4E²¹. Perhaps the presence of scaffold proteins on ribosomes is a widespread tool for tuning kinase activity on selected ribosomal populations. □

Methods

Cell lines, immunofluorescence and mutagenesis

COS-7, HeLa Tet-off and human A431 cells were cultured in DMEM according to standard procedures. TPA (Calbiochem) treatment was performed at 50–100 nM. The following antibodies were used: rabbit anti-eIF6⁵, mouse anti-RACK1 IgM monoclonal antibody (Transduction Laboratories), mouse anti-lamin B (Santa Cruz), monoclonal antibodies against PKC isoforms (Transduction Laboratories). Antibodies against ribosomal markers, originally by T. Sato, were prepared in our laboratory as detailed in ref. 22 by Primm Biotech. Secondary antibodies were from DAKO. Immunofluorescence was performed as previously described⁸. Site-directed mutagenesis was performed according to the QuikChange XL site-directed mutagenesis kit (Stratagene).

Polysome analysis and fractionation

Ribosome preparation and analysis were essentially performed as previously described⁸. Profiles for Figs 1 and 2 were recorded on a chart. All the others were digitally recorded. For protein analysis, collected fractions were precipitated with trichloroacetic acid (TCA), run on SDS–PAGE (polyacrylamide gel electrophoresis) and blotted. Nuclear–cytoplasmic fractionation was performed as in ref. 23 and controlled as described in Supplementary Information.

Yeast two-hybrid screening

The full-length human eIF6 clone was used in the screening of a HeLa cell library, and validation studies²⁴ were performed according to the GAL4 system (Supplementary Methods).

Extraction, immunoprecipitation and western blotting

Cells were scraped in modified RIPA buffer (10 mM Tris–HCl pH 7.2, 0.5% deoxycholic acid, 0.05% SDS, 150 mM NaCl, 1 mM EDTA and 1 mM PMSF). The lysate was incubated for 30 min at 4 °C and cleared by centrifugation. The amount of protein recovered was measured by the bicinchoninic acid protein assay (Pierce). For co-immunoprecipitation assays, 3 mg of extracts were pre-cleared with protein A beads (Pharmacia) and centrifuged at 2,000g for 5 min at 4 °C. Supernatants (1.5 mg proteins) were incubated for 2 h at 4 °C with either anti-eIF6 antibody or with pre-immune rabbit sera. Protein A beads were added and incubated for 1 h at 4 °C. The beads were washed four times with 1% Triton X-100 in PBS and resuspended in Laemmli buffer. Immunoprecipitates were resolved on SDS–PAGE, transferred to Immobilon-P membranes (Millipore) and probed by western blotting. Western blotting for transfected proteins was performed using either anti-eIF6 antibodies or anti-haemagglutinin (Monoclonal antibody 12CA5, Roche Diagnostics).

Pull-down assay and far-western blotting

A431 lysates were prepared as for immunoprecipitation, pre-cleared with amylose beads, and incubated for 3 h at 4 °C with 5 µg of MBP or 2 µg of MBP–RACK1 bound to amylose beads. Beads were pelleted and washed four times. After resuspension in Laemmli buffer, the supernatants were separated using SDS–PAGE and blotted with the relevant antibodies.

For far-western studies, the MBP–RACK1 fusion protein was cleaved with Factor Xa, purified, run on native PAGE gel, and transferred to Immobilon P-membranes that were incubated with biotin-labelled eIF6. Bound protein was revealed using peroxidase-conjugated streptavidin (Amersham) and ECL.

Production of recombinant proteins

Recombinant RACK1 fused to MBP was purified from *Escherichia coli* according to the manufacturer's protocol (New England Biolabs). Untagged, recombinant forms of eIF6 were subcloned into pET23-a vectors, expressed in *E. coli* and purified to homogeneity (Supplementary Methods).

OligodT pull-down

Cells were washed and lysed at room temperature in hypotonic buffer as described in ref. 25. Centrifugation (25,000g) was performed at room temperature for 3 min. Lysates corresponding to 1 million cells were incubated with oligodT tablets (Invitrogen) for 1 h at room temperature. After incubation, mixtures were pelleted, washed three times, resuspended in Laemmli buffer and analysed as described elsewhere.

Transfections and measurement of translation

COS cells or Tet-off HeLa cells were used for these experiments. For experiments with the luciferase reporter, the Tet-off system was used²⁶ and cells were transfected with calcium phosphate or lipofectamine. COS cells were used for metabolic-labelling experiments. Transfection was performed with lipofectamine with a required transfection efficiency of at least 85% (Supplementary Methods).

Reticulocyte experiments

Reticulocyte extracts (Promega) of pure ribosomes were pre-incubated for 30 min at 30 °C with the indicated amount of purified factors. After preincubation, luciferase mRNA was

added to the mix and incubated for 30 min at 30 °C. The reaction was stopped, and luciferase activity was measured using a luminometer.

In vitro phosphorylation assay by PKC

Standard procedures with various modifications were used and are detailed in Supplementary Methods. For rescue experiments, 6 µg of freshly prepared eIF6 was used for each sample. To measure eIF6 phosphorylation, 1 µg of recombinant protein was used.

[³²P]orthophosphate labelling of COS7 cells

COS7 cells were used for these experiments. Metabolic labelling was performed with carrier-free [³²P]orthophosphate (Amersham) after phosphate starving and according to the conditions specified in Supplementary Methods. Labelling was performed with 700–2,000 µCi ml^{−1} for at least 4 h.

Subunit joining assay

Salt-washed ribosomal subunits, a gift from T. Pestova, were used²⁷. The assay, based on increasing Mg²⁺ concentration to induce 80S formation¹¹, was slightly modified to include recombinant factors.

Received 22 August; accepted 10 October 2003; doi:10.1038/nature02160.

1. Pestova, T. V. *et al.* Molecular mechanisms of translation initiation in eukaryotes. *Proc. Natl Acad. Sci. USA* **98**, 7029–7036 (2001).
2. Gingras, A. C., Raught, B. & Sonenberg, N. eIF4 initiation factors: effectors of mRNA recruitment to ribosomes and regulators of translation. *Annu. Rev. Biochem.* **68**, 913–963 (1999).
3. Rhoads, R. E. Signal transduction pathways that regulate eukaryotic protein synthesis. *J. Biol. Chem.* **274**, 30337–30340 (1999).
4. Harding, H. P. *et al.* Regulated translation initiation controls stress-induced gene expression in mammalian cells. *Mol. Cell* **6**, 1099–1108 (2000).
5. Biffo, S. *et al.* Isolation of a novel β4 integrin-binding protein (p27^{BBP}) highly expressed in epithelial cells. *J. Biol. Chem.* **272**, 30314–30321 (1997).
6. Si, K., Chaudhuri, J., Chevesich, J. & Maitra, U. Molecular cloning and functional expression of a human cDNA encoding translation initiation factor 6. *Proc. Natl Acad. Sci. USA* **94**, 14285–14290 (1997).
7. Ron, D. *et al.* Cloning of an intracellular receptor for protein kinase C: a homolog of the β subunit of G proteins. *Proc. Natl Acad. Sci. USA* **91**, 839–843 (1994).
8. Sanvito, F. *et al.* The β4 integrin interactor p27(BBP/eIF6) is an essential nuclear matrix protein involved in 60S ribosomal subunit assembly. *J. Cell Biol.* **144**, 823–837 (1999).
9. Si, K. & Maitra, U. The *Saccharomyces cerevisiae* homologue of mammalian translation initiation factor 6 does not function as a translation initiation factor. *Mol. Cell. Biol.* **19**, 1416–1426 (1999).
10. Wood, L. C., Ashby, M. N., Grunfeld, C. & Feingold, K. R. Cloning of murine translation initiation factor 6 and functional analysis of the homologous sequence YPR016c in *Saccharomyces cerevisiae*. *J. Biol. Chem.* **274**, 11653–11659 (1999).
11. Valenzuela, D. M., Chaudhuri, A. & Maitra, U. Eukaryotic ribosomal subunit anti-association activity of calf liver is contained in a single polypeptide chain protein of *M_r* = 25,500 (eukaryotic initiation factor 6). *J. Biol. Chem.* **257**, 7712–7719 (1982).
12. Mochly-Rosen, D., Smith, B. L., Chen, C. H., Disatnik, M. H. & Ron, D. Interaction of protein kinase C with RACK1, a receptor for activated C-kinase: a role in β protein kinase C mediated signal transduction. *Biochem. Soc. Trans.* **23**, 596–600 (1995).
13. Rodriguez, M. M., Ron, D., Touhara, K., Chen, C. H. & Mochly-Rosen, D. RACK1, a protein kinase C anchoring protein, coordinates the binding of activated protein kinase C and select pleckstrin homology domains *in vitro*. *Biochemistry* **38**, 13787–13794 (1999).
14. Ron, D. *et al.* Coordinated movement of RACK1 with activated βIIIPKC. *J. Biol. Chem.* **274**, 27039–27046 (1999).
15. Schechtman, D. & Mochly-Rosen, D. Adaptor proteins in protein kinase C-mediated signal transduction. *Oncogene* **20**, 6339–6347 (2001).
16. Groft, C. M., Beckmann, R., Sali, A. & Burley, S. K. Crystal structures of ribosome anti-association factor IF6. *Nature Struct. Biol.* **7**, 1156–1164 (2000).
17. Senger, B. *et al.* The nucle(ol)ar Tifp and Eflp are required for a late cytoplasmic step of ribosome synthesis. *Mol. Cell* **8**, 1363–1373 (2001).
18. Chantrel, Y., Gaisne, M., Lions, C. & Verdiere, J. The transcriptional regulator Hap1p (Cyp1p) is essential for anaerobic or heme-deficient growth of *Saccharomyces cerevisiae*: Genetic and molecular characterization of an extragenic suppressor that encodes a WD repeat protein. *Genetics* **148**, 559–569 (1998).
19. Proud, C. G. *et al.* Interplay between insulin and nutrients in the regulation of translation factors. *Biochem. Soc. Trans.* **29**, 541–547 (2001).
20. Peterson, R. T. & Schreiber, S. L. Translation control: connecting mitogens and the ribosome. *Curr. Biol.* **8**, 248–250 (1998).
21. Pyronnet, S. *et al.* Human eukaryotic translation initiation factor 4G (eIF4G) recruits mnk1 to phosphorylate eIF4E. *EMBO J.* **18**, 270–279 (1999).
22. Nadano, D. *et al.* Preparation and characterization of antibodies against human ribosomal proteins: heterogeneous expression of S11 and S30 in a panel of human cancer cell lines. *Jpn. J. Cancer Res.* **9**, 802–810 (2000).
23. Spratt, S. C., Hammond, K. D. & Savage, N. Subcellular fractionation of murine erythroleukemic cells: distribution of protein kinases. *Anal. Biochem.* **194**, 407–412 (1991).
24. Estojak, J., Brent, R. & Golemis, E. A. Correlation of two-hybrid affinity data with *in vitro* measurements. *Mol. Cell. Biol.* **15**, 5820–5829 (1995).
25. Ho, J. H., Kallstrom, G. & Johnson, A. W. Nascent 60S ribosomal subunits enter the free pool bound by Nmd3p. *RNA* **6**, 1625–1634 (2000).
26. Baron, U. & Bujard, H. Tet repressor-based system for regulated gene expression in eukaryotic cells: principles and advances. *Methods Enzymol.* **327**, 401–421 (2000).
27. Pestova, T. V. *et al.* The joining of ribosomal subunits in eukaryotes requires eIF5B. *Nature* **403**, 332–335 (2000).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank E. Rizzo, S. Tognin, A. Donadini, A. M. Barbieri and L. Castelli for technical advice; for providing salt-washed purified ribosomal subunits, we thank T. Pestova. For sharing reagents, information and critical advice, we thank A. Bachi, T.-A. Sato, F. Loreni, F. Amaldi, P. Linder, J. Traugh, A. Johnson, A. Hinnebusch, A. Burlando, L. Spemulli, E. Villa, J. Verdieri, D. Ron, T. Dever, C. Groft and M. Foiani. This work was supported by AIRC, MURST (P.C.M., S.B.). This study was carried out under the framework of the Italian MUIR Center of Excellence in Physiopathology of Cell Differentiation.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to S.B. (biffo.stefano@hsr.it).

corrigenda

Colorectal carcinomas in mice lacking the catalytic subunit of PI(3)K γ

Takehiko Sasaki, Junko Irie-Sasaki, Yasuo Horie, Kurt Bachmaier, Jimmie E. Fata, Martin Li, Akira Suzuki, Dennis Bouchard, Alexandra Ho, Mark Redston, Steven Gallinger, Rama Khokha, Tak W. Mak, Phillip T. Hawkins, Len Stephens, Stephen W. Scherer, Ming Tsao & Josef M. Penninger

Nature 406, 897–902 (2000).

In this Letter, we reported that mice lacking the catalytic subunit of PI(3)K γ on 129J backgrounds develop invasive colorectal tumours. We have observed this cancer phenotype in mouse lines derived from two different targeted embryonic stem-cell clones for at least two years. However, after backcrossing these mice onto a C57BL/6 background, we now find that the tumour phenotype has disappeared. Also, when we retargeted the allele in different ES cells using the same targeting construct, no tumours developed. Inactivation of PI(3)K γ in mice therefore does not in itself cause colon cancer, but may require additional factors—for example, the impaired immunity of p110 $\gamma^{-/-}$ mice may make them susceptible to tumours triggered by environmental and infectious agents¹. Our finding that overexpression of p110 γ in different human colon-cancer cells results in decreased cell growth still stands, as does the frequent downregulation of p110 γ protein in colon-cancer patients (since independently confirmed²).

Some authors (Y.H., A.H., M.R., S.G. and T.W.M.) are not co-signatories to this statement. □

1. Erdman, S. E. *et al.* CD4⁺ CD25⁺ regulatory T lymphocytes inhibit microbially induced colon cancer in Rag2-deficient mice. *Am. J. Pathol.* **162**, 691–702 (2003).
2. Semba, S. *et al.* Down-regulation of PIK3CG, a catalytic subunit of phosphatidylinositol 3-OH kinase, by CpG hypermethylation in human colorectal carcinoma. *Clin. Cancer Res.* **8**, 3824–3831 (2002).

Cyanophages infecting the oceanic cyanobacterium *Prochlorococcus*

Matthew B. Sullivan, John B. Waterbury & Sallie W. Chisholm

Nature 424, 1047–1051 (2003).

In this Letter, the data for host strains WH8012 and WH8109 in Fig. 1 were inadvertently reversed. This mistake does not affect any of the conclusions made in the paper. □

Zero thermal expansion in YbGaGe due to an electronic valence transition

James R. Salvador, Fu Guo, Tim Hogan & Mercouri G. Kanatzidis

Nature 425, 702–705 (2003).

In this Letter, both the legend to Fig. 1 and the 'X-ray diffraction data' section of the Methods contain errors in the crystallographic data. In the Methods, for both YbGaGe and YbGaSn, $Z = 4$ (not 3). Hence, for YbGaGe, density $d_{\text{calc}} = 8.149 \text{ g cm}^{-3}$ and absorption coefficient $\mu = 57.752 \text{ mm}^{-1}$. For YbGaSn, $d_{\text{calc}} = 8.151 \text{ g cm}^{-3}$ and $\mu = 48.612 \text{ mm}^{-1}$. In the legend to Fig. 1, the atomic coordinates ($\times 10^4$) should be as follows: In both compounds, for Yb(1) (x, y, z) = (0, 0, 1/4). For Yb(2) (x, y, z) = (0, 0, 0). In YbGaGe, for Ga (x, y, z) = (1/3, 2/3, 0.1532) and for Ge (x, y, z) = (1/3, 2/3, 0.6128). In YbGaSn, for Ga (x, y, z) = (1/3, 2/3, 0.1634) and for Sn (x, y, z) = (1/3, 2/3, 0.6146). In addition, line 21 of the left column of page 704 should have read "...with one showing NTE and one PTE." □

Cdc6 cooperates with Sic1 and Hct1 to inactivate mitotic cyclin-dependent kinases

Arturo Calzada, Maria Sacristán, Elisa Sánchez & Avelino Bueno

Nature 412, 355–358 (2001).

In this Letter it was reported that an amino-terminal CDK-inhibitory domain of the Cdc6 replication protein was required for efficient mitotic exit, and that removal of this domain lethally blocked mitotic exit in cells lacking the Sic1 CDK inhibitor, demonstrating essential CDK-inhibitory control of CDK activity. However, Archambault *et al.*¹ have shown that viable mitotic-exit-competent strains lacking both CDK-inhibitors can be constructed, a finding that has also been confirmed by the authors. Archambault *et al.*¹ suggest that CDK-inhibitors are not required for mitotic exit in the wild-type cell cycle. However, A.C. *et al.* (manuscript in preparation) have evidence that strains lacking both CDK inhibitors may play a non-essential role in the regulation of mitotic exit. This point is still under investigation. □

1. Archambault, V. *et al.* Genetic and biochemical evaluation of the importance of Cdc6 in regulating mitotic exit. *Mol. Biol. Cell* (in the press); advance online publication 5 September 2003 (doi:10.1091/mbc.E03-06-0384).

Sustained division of the attentional spotlight

M. M. Müller, P. Malinowski, T. Gruber & S. A. Hillyard

Nature 424, 309–312 (2003).

The support of the University of Liverpool, where some of the work was carried out (by M.M.M., P.M. and T.G.), is acknowledged. □

Go for the soft cell

For cell biologists and those with similar callings.

Alexa-Fluoronanogold

Nanoprobes www.nanoprobes.com

One probe, two answers

The Alexa-Fluoronanogold product line combines the brightness of Alexa fluorophores with 1.4 nm Nanogold in one probe, permitting visualization of targets by fluorescence in a light microscope (LM) followed by more detailed ultrastructural study in an electron microscope (EM). Thus researchers can save time by processing samples for EM only after acceptable labelling is confirmed at the LM level. Since the two reporters are on a single Fab' antibody probe, the two answers may be directly correlated. Fab' antibody fragments are used to ensure easy penetration into tissues and high-resolution labelling. Alexa dyes are brighter and bleach less than fluorescein, improving the fluorescent signal. Covalent attachment of both the Alexa dye and Nanogold to the Fab' antibody fragment ensure stability and shelf life.

SZX7 stereo microscope

Olympus www.olympus-europa.com

A zoom with a view

The Olympus SZX7 incorporates advanced infinity-corrected optics and modular design, making it a flexible system for different applications and for upgrading. The



Olympus SZX7: to infinity and beyond.

zoom ratio is 7:1, covering the magnification range from 8× to 56× (using 1× objective/10× eyepiece). A choice of observation tubes includes a 5–45° tilting binocular tube for optimum comfort. Long working distance objectives and intermediate application tubes are available, as well as a variety of illumination systems. Resolution is enhanced to 600 lines per mm and the precision optics deliver an exceptionally natural colour together with flat, distortion-free images right to the edge of the field of view. Observation methods include GFP fluorescence and oblique illumination. The click zoom feature means that an observation position can be easily reproduced, which is particularly useful for photographic recording.



PET sound: INTEGRA's disposable bioreactor.

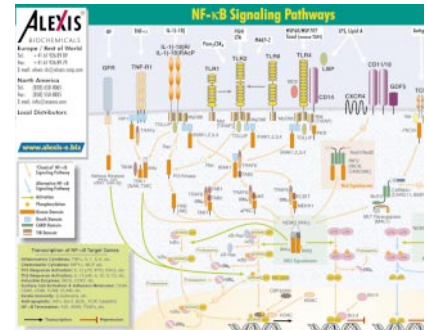
CELLine adhere

INTEGRA Biosciences

www.integra-biosciences.com

A throwaway line

The CELLine adhere is a novel, disposable flask bioreactor designed for long-term, high-level protein expression from adherent cells. The reactor incorporates a woven polyethylene terephthalate (PET) matrix in the cell compartment to provide an ideal surface for attachment of anchorage-dependent cells such as HEK, BHK and CHO. The system uses novel membrane technology to separate the cultivation chamber, with an upper semi-permeable membrane through which nutrient can diffuse and a lower one that allows diffusion of gases. Separate ports allow selective access to the upper nutrient supply chamber and the central cultivation chamber. This compartmentalized arrangement means that medium can be exchanged without influencing the function or growth of the cells. With achievable cell densities of between 10^7 and 10^8 cells per ml, CELLine adhere provides cell concentrations typically two orders of magnitude higher than conventional static cell-culture techniques.



Free wall: lab decor from Alexis.

NF-κB wall chart

Alexis www.alexis-e-biz

NF-κB signalling pathways at a glance

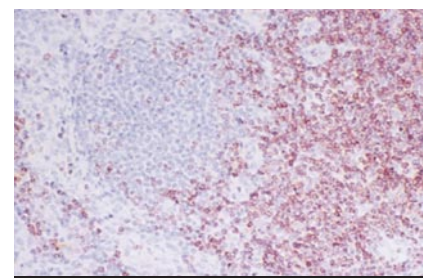
This full-colour wall chart from Alexis shows the relationships between the various components in the NF-κB signalling pathways as well as associations between pathways, and includes many recently identified enzymes and substrates. The chart is organized by component locations within the cell and by function. It is a useful educational and quick reference tool for anyone studying or working in this area of research. The chart can be downloaded from the company's web site in JPEG format and is also available in printed form.

Polyclonal CD3 antibody

Zymed Laboratories www.zymed.com

Rabbit anti-CD3 concentrate

Human CD3 is a glycoprotein consisting of five polypeptide chains (gp20, gp26, p16, p20 and p28). The CD3 complex is highly involved in T-cell activation and signal transduction. CD3 antigens are T-cell-specific and present in resting and activated T lymphocytes. Zymed's polyclonal CD3 antibody has been characterized for immunohistostaining of frozen and formalin-fixed, paraffin-embedded tissue sections. It is used in the differentiation of precursor T-cell acute lymphoblastic lymphoma from its B-cell counterpart.



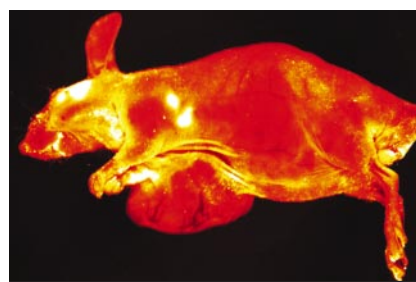
CD3 staining: zero in on T cells.

ClonePix

Genetix www.genetix.com

Easy pickings

The ClonePix automated mammalian cell picker from Genetix will image, identify and select hybridoma cell colonies from either semi-solid media or monolayers growing on cell-culture dishes and deposit them into microplates for subculture. ClonePix enables a laboratory to screen over 5,000 colonies per week, a fivefold increase over manual picking. It can also be used to automate the selection of transfected cells and cell lines for pharmaceutical screening. The system is fully enclosed, with HEPA filtration and positive pressure to minimize external contamination. A built-in ultraviolet lamp can be used to sterilize the working area, reducing the risk of contamination through manual intervention.



The whole picture: Kodak's new camera in action.

IS2000MM

Kodak www.kodak.com

Molecular imaging in vivo

Kodak's IS2000MM (Image Station 2000 Multi-Modal Imaging System) features cooled CCD camera technology for fluorescence, luminescence, radiographic and chromogenic imaging in life science research. Multi-wavelength illumination improves detection sensitivity and reduces background for most fluorescence labels, and allows imaging of multi-fluorochrome assays. In addition, the ability to select and apply longer excitation wavelengths (green to near-infrared) improves the penetration of light into tissue, allowing whole-body *in vivo* molecular imaging via fluorescently tagged biomolecules in

living animals and plants. The system also provides quantitative imaging of luminescent, fluorescent, radiographic and colourimetric labelled biomolecules in a variety of common assay formats, including membranes, electrophoresis gels, plates, tissues, and *in vivo* assays.

Antibodies for membrane trafficking

Covance Research Products www.crpinc.com

Special agents

Covance manufactures several specialized antibodies for intracellular and extracellular markers. They include monoclonal antibodies that recognize specific organelle proteins such as mannosidase II and nuclear pore complex proteins. These antibodies have been shown to be useful in understanding the roles and relationships of proteins in the Golgi apparatus and nuclear envelope. A polyclonal antibody specific for involucrin is also available, and can be used to study terminal differentiation in the upper layers of the epidermis. All antibodies can be used in immunofluorescence studies and have very high titres.

Bovine serum albumins

Amresco www.amresco-inc.com

Cost-effective solutions for molecular biology

Amresco has added three new products to its line of bovine serum albumin (BSA) solutions. Bovine serum albumin 20 mg per ml is a solution of BSA (albumin, bovine, fraction V — cold alcohol isolation) for molecular biology research. It contains 0.1% sodium azide and is prepared in deionized water. This solution is free from RNase, DNase and protease, and remains stable for two years. Bovine serum albumin, 20% and 30% consist of concentrated 20% and 30% BSA in 0.85% sodium chloride, also containing 0.1% sodium azide. They are cost-effective and are aseptically filled in order to provide consistent detection results.

HcRed1 localization vectors

BD Biosciences Clontech www.clontech.com

Far-red fluorescence

BD Biosciences Clontech has expanded its BD Living Colors range of subcellular localization vectors to include pHcRed1-Nuc and pHcRed1-Mito. These vectors express the far-red fluorescent protein HcRed1 fused to a localization tag. Because of its far-red shifted fluorescence, HcRed1 is easily distinguished from the company's other fluorescent proteins. pHcRed1-Nuc expresses HcRed1 with a C-terminal nuclear localization signal that localizes the chromophore in the cell nucleus. pHcRed1-Mito expresses HcRed1 with an N-terminal mitochondrial targeting sequence that targets the chromophore to the mitochondria. Whereas conventional immunostaining requires fixing and washing of cells several times, these subcellular



In the red: local knowledge from Clontech.

localization vectors allow direct, non-invasive visualization of living cells.

Dri-Pure vacuum ramp

Genevac www.genevac.co.uk

For a smooth run

Rapid drying with no loss of sample and no cross-contamination due to solvent bumping are the benefits claimed by Genevac for Dri-Pure technology, designed for use with the company's centrifugal solvent evaporators. The anti-bumping feature works by controlling the reducing pressure in the evaporation chamber, in combination with an increase in rotor speed to achieve greater *g* force. In addition, heat flow to the samples must be reduced by switching off the infrared heat lamps during the pressure ramping stage. This results in smooth, controlled solvent removal within the optimum time-span, a benefit particularly useful for chemists working in parallel synthesis, drug discovery and compound purification. Dri-Pure technology is available for all current models of Genevac's centrifugal evaporators, from the EZ-2 and DD-4X benchtop units to the production-scale Mega 1200.

ScanView

Image Solutions www.imsol.co.uk

Rare cells in focus

The ScanView system for automatic rare cell detection features a motorized XY stage, automated microscope focus control and Z-axis. The XY stage will relocate automatically to the selected cell coordinates from the database. It will also register the coordinates of an acquired cell automatically, enabling users to return to previously acquired cells for visual inspection or relocate to cells that were found by another scanning system on the network. Automated focus control and Z-axis ensure that both live views and acquired images are of the best quality and delivers automatic two-dimensional imaging from an image stack. Scanning along the Z-axis produces a set of images in different focal planes, allowing export of three-dimensional data. The system includes a colour camera with acquisition software that allows users to receive a one-shot qualitative colour image as seen with visual inspection.

These notes are compiled in the Nature office from information provided by the manufacturers.

Contacts

Publisher: Ben Crowe

Editor: Paul Smaglik

Marketing Manager: David Bowen

European Head Office, London

The Macmillan Building
4 Crinan Street
London N1 9XW, UK
Tel +44 (0) 20 7843 4961
Fax +44 (0) 20 7843 4996
e-mail: naturejobs@nature.com

Naturejobs Sales Director:
Nevin Bayoumi (4978)

UK/ RoW/ Ireland:

Matt Powell (4953)

Andy Douglas (4975)

Frank Phelan (4944)

Netherlands/ Italy/ Spain/

Portugal/ Belgium:

Evelina Rubio Hakansson (4973)

Scandinavia: Sille Opstrup (4994)

France/ Switzerland:

Amelie Pequignot (4974)

Natureevents:

Paul Constant (4954)

Production Manager:

 Billie Franklin

To send materials use London
address above.

Tel +44 (0) 20 7843 4814

Fax +44 (0) 20 7843 4996

e-mail: naturejobs@nature.com

Naturejobs web development:

Tom Hancock

Naturejobs online production:

Ben Lund

European Satellite Office

Germany/ Austria:

Patrick Phelan, Odo Wulffen

Tel + 49 89 54 90 57 11/2

Fax + 49 89 54 90 57 20

e-mail: p.phelan@nature.com

o.wulffen@nature.com

US Head Office, New York

345 Park Avenue South,

10th Floor, New York, NY 10010-1707

Tel +1 800 989 7718

Fax +1 800 989 7103

e-mail: naturejobs@natureny.com

US Sales Manager:

 Peter Bless

US Advertising Coordinator:

Linda Adam

Japan Head Office, Tokyo

MG Ichigaya Building (5F),

19-1 Harakatsutamachi,

Shinjuku-ku,

Tokyo 162-0841

Tel +81 3 3267 8751

Fax +81 3 3267 8746

Asia-Pacific Sales Director:

Rinoko Asami

e-mail: rasami@naturejpn.com

naturejobs

Making choices

Two days of Naturejobs career talks at the International Biotech meeting last week in London yielded different career stories — but several common denominators. Some of the speakers had left the lab bench because they didn't want to focus too tightly. Ann Gimalac, Gail Cardew and Ann Van Gysel took communication-based jobs instead — Gimalac for Genset/Serono outside Paris, Cardew for the Royal Institution in London and Van Gysel for the Flanders Interuniversity Institute for Biotechnology. In addition to a broad curiosity, Otello Stampacchia confessed to a short attention span and a difficult time mastering basic lab techniques. So he networked himself into the venture-capital business at NIB Capital, a private company in Geneva, Switzerland.

Their scientific background was the foundation for making the switch. But soft skills made it possible. When scientists are telling other people what they do, says Stampacchia, they should keep it short, sweet and memorable. Scientists should also follow that advice when assembling a resumé, says Hannah Thompson, a manager with the Oxford scientific recruitment agency SRG.

The route from bench to 'alternative' career is considered to be easier in industry than in academia. Jean-Erick Ancel traded in ten years of making molecules at Rhône-Poulenc for a career in marketing. The skills and industry knowledge he had accumulated made the switch a natural one. David Reynolds, a research fellow at Merck Sharpe Dohme, says that having this inside knowledge is a good way to move from the bench into human resources or communications, as well as into marketing.

The hardest choice may be deciding what to do. Claude Feuerstein, former president of Joseph Fourier University in Grenoble, France, says "Do what you want". Getting that right makes obstacles more bearable — and success more satisfying.

Paul Smaglik
Naturejobs editor



Contents

CAREER VIEW

Nuts & Bolts

Get organized

Young Scientist

Women as mentors

Movers

Alex Matter

p588

WWW.NATUREJOBS.COM

Career centre

Information on the
scientific job market

FOCUS

SPOTLIGHT

RECRUITMENT

ANNOUNCEMENTS

EVENTS

NUTS & BOLTS

YOUNG SCIENTIST

Women mentoring women

Scientists generally agree that finding mentors for young people attracts them to, and keeps them in, science. This is particularly true for women and minorities. But with few women and minorities at faculty level, where should the mentors come from? This question spurred the development of a mentoring programme at Caltech in 2003 called Women Mentoring Women (WMW).

A cooperative effort of several groups on campus, led by the women's centre, the WMW programme pairs graduate students with female postdocs. It also offers group mentoring in the form of book discussions about women in science and professional-development presentations given by faculty members.

This spring, 22 graduate women and postdocs participated in the pilot WMW. The evaluations showed that postdoc mentors gained as much as the graduate students, reinforcing the belief that mentoring benefits both sides.

In the current academic year, WMW's numbers have risen to 62, and it is expected to expand even further in 2004–05. Its participants hope it will serve as a model that will spread to other institutions suffering a dearth of female faculty members, providing mentors for young women scientists at a crucial point in their careers. ■

Helen McBride, Caltech Postdoctoral Association

Scientists are among the most educated people on the planet, yet many fail to use effective employment strategies. Uncovering opportunities and jockeying for position requires more street sense than academic ability. So focus on your goals and use the following guidelines.

Take a multidimensional approach. Move beyond passive job-seeking to more active methods: face-to-face networking, searching the Internet, checking scientific-association journals and getting to know recruiters. For the best use of your time, try all of these from the start.

Keep it live. Spend most of your time networking or meeting potential employers in person. The Internet offers job postings and employer information, but the decision to make an offer rests with human beings. At a career fair one drug-company recruiter said his firm received more than 1,000 applications a



With Deb Koen
Careers consultant

week through its website. A referral by an employee gives you much better odds of landing an interview. That's equally true in academia, clinical settings and small enterprises.

Structure your time. Schedule time each day for your search. In the initial preparatory phase (self-assessment, employer research, CV/resumé development) you can work day or night. But the next, contacting phase consists of phone calls and meetings that are best held during the working day.

Go the extra mile. In a tightened economy, most

job seekers are covering basic etiquette. To stand out, follow up every networking call and job interview with a personal letter or e-mail, reinforcing your interest and highlighting your potential contributions.

Work on your image. Apart from your scientific expertise, nothing carries more weight in winning you a job offer than the way you present yourself in writing and in person. Begin with a polished CV or resumé. Then develop speaking points about your goals and accomplishments that convey who you are and what you have to offer. Practice these with a colleague you can count on for honest criticism. After you've made selected employers aware of your background and skills, it's your communication that will close the deal. ■

Deb Koen is vice-president of Career Development Services and a columnist for The Wall Street Journal's CareerJournal.com.

MOVERS

Alex Matter, Director, Novartis Institute for Tropical Diseases, Singapore



When Alex Matter went into industry 28 years ago, academic colleagues warned that the move would mark "the end of my scientific career", he says. Since then he has managed to usher a successful cancer drug through the development pipeline and has now moved back into a more academic environment.

That's not to say his career path was easy when it veered off into industry. He found the shift initially "pretty horrible",

because many of the people who had recruited him were gone by the time he arrived, making it hard for him to get his bearings. "I had to find my way through this labyrinth," Matter says. After a year of negotiating the maze, he assembled his research team — only to learn that the project they wanted to take on was to be killed. "I discovered this sort of thing is very routine," Matter says.

His advice to anyone considering a similar path in drug discovery? Get used to adversity. Having projects shelved, stalled or killed "should not deter you", he says. It taught him to find bosses who were on the same page. That strategy led him to Ciba-Geigy, where a friend said, "Let's do something in cancer".

That friend left not long after Matter joined, too. But Matter spent the next three months in the library, inspired by a string of oncogene discoveries. He was convinced that there must be a better way to target cancer: chemotherapy, he

felt, had "plateaued" and immunotherapy seemed ineffective. So he became captivated by kinases and hit the books.

The gap between inspiration and results proved long and tedious. He wrote down the general concept underlying Glivec/Gleevec (imatinib mesylate) in 1983, formed a team that had a molecule ready in 1992, and after many ups and downs in preclinical studies, it finally entered clinical trials in 1998. The drug was approved in 2001.

That slog also taught him something else about surviving in the high-stakes world of drug discovery. "You have to have bread-and-butter projects in order to pursue the high-risk stuff," he says.

Now, he thinks his goal at the Novartis Institute of finding drug leads for tuberculosis and dengue fever within three years is "immensely do-able, as we have much more information than when we started on Glivec". He has also learned that tenacity reaps rewards. ■

CV **1996–2003:** Head of the Oncology Therapeutic Area at the Novartis research department in Basel, Switzerland
1983–1996: Head of the Cancer and Bone/Bone Metabolism Therapeutic Area of Ciba-Geigy in Basel
1986–1997: Head of AIDS research at Ciba-Geigy
1980–83: Director of the Immunology Research Laboratories of Schering in France
1975–1980: Group leader in tumour immunology at F. Hoffmann-La Roche in Basel